Birdsong as a model for human language: controversies and further steps

Author: Daniëlle Fluks (S2893207)
Supervisor: prof. dr. S. Verhulst
Date: February 2019
Research course: Gedragsbiologie research
Bachelor thesis

Abstract

Humans are often thought to be unique in their linguistic abilities. Yet, the idea that certain features of the human language system might be found in the animal kingdom as well goes at least as far back as Darwin. More recently, behavioural, neuroanatomical and genetic similarities between the human language system and several other species were found, most notably in songbirds. This has led to a discussion on the suitability of birdsong as a model of human language. Such a model would be quite helpful, as the complexity of the human language, as well as practical and ethical limitations, make it difficult to study the human language system on a genetic and neurological level.

Unlike human language, birdsong has not been shown to contain meaningful units, thus supposedly lacking any form of semantics. Yet, on the phonological and syntactical level, as well as with regards to vocal learning, the similarities between birdsong and human language appear to be striking, though they are not uncontroversial. This has led to a discussion between scientists that think birdsong could be a good model of human language, and scientists that do not think birdsong is complex enough to fulfil this role.

More research into the specific similarities between birdsong and human language are necessary to determine if and in what way birdsong could contribute to our understanding of the human language system. Furthermore, the current focus on Universal Grammar as the biological basis of language might impair our understanding of the relationship between birdsong and human language, as it prevents alternative views of a biological basis of these systems from being considered in the discussion.

Contents

Introduction ... 3
1 - Semantics ... 5
2 - Phonology and prosody ... 7
2.1 - Similarities between bird and human phonology and prosody ... 7
2.2 - Phonological development ... 8
2.3 - Prosody ... 9
3 - Syntax ... 10
3.1 - Artificial grammar studies ... 10
3.2 - Syntactic complexity ... 11
4 - Vocal learning ... 13
4.1 - The critical period ... 13
4.2 - Learning biases ... 14
5 - Discussion ... 15
5.1 - General discussion ... 15
5.2 - Semantics ... 16
5.3 - Phonology ... 17
5.4 - Syntax ... 17
5.5 - Vocal learning ... 18
References ... 20

Introduction

The thought of bird vocalisations being similar to human language is nothing new. Darwin (1871/2016) already linked the two phenomena, calling birdsong the “nearest analogy to language.” Modern advances in genetics and neurobiology have suggested that this statement is not just true on a behavioural level, but on a more fundamental level as well (e.g. Condro & White, 2014; Petkov & Jarvis, 2012; Pfenning et al., 2014). In the last decade, research has become focused on birdsong as a possible model of human speech. In particular, researchers are interested in the ways the system of human language and the system of birdsong share features, both mechanistically and behaviourally, thus determining the limits of birdsong as a model.

But why is such a model important to begin with? Research on human language is mostly conducted on human subjects within the discipline of linguistics. This limits the possible types of experiments. For example, it is very difficult to study the genetics and pathways of language in humans, as it is usually not possible to simply block a part of the pathway or switch off a gene. Because of this, our knowledge about the biological compound of language remains limited. The work of linguistics could be described as a mainly top-down approach; simple rules are interfered from meticulously studying human language utterings. Biology could add a more bottom-up approach, studying the pathways involved with language, given a suitable model species. A model species could be used to extensively investigate these specific features, mapping out their pathways and ultimately getting to an understanding of this basic form of the language system. This could then be linked to the knowledge linguistics has generated on human language, which is presumed to be more complex. However, it has to be taken in mind that such a comparison can only be made if the pathways show a high level of either conservation or convergent evolution between humans and the model species.

The idea of a model species for human language has gained increased interest over the last decade. Songbirds are the prime candidate. These birds are said to show vocal learning in a way very similar to vocal learning in humans (Tchernichovski & Marcus, 2014). Furthermore, the syntactical and phonological aspects of birdsong closely resemble human speech as well (Ten Cate, 2014). Genetic and neurological work has suggested that the pathways might show a relatively high level of convergent evolution between birdsong and human speech (Pfenning et al., 2014). This is especially remarkable because of the great phylogenetic distance between humans and songbirds. Evolutionary closer species, like non-human primates, do not seem to show the same level of complexity in their vocal communication, for example by not clearly displaying a syntactical structure in their vocalisations (Petkov & Jarvis, 2012). However, the idea of songbirds as a model species for human language does not remain uncontroversial and has led to a clear-cut debate. On one

side there is a group of researchers, including Carel ten Cate and Toshitaka Suzuki, that conducts experimental work on the similarities between birdsong and features of human language, and believes songbirds would be a good model species. On the other side we find a group, including Johan Bolhuis and Robert Berwick, that believes birdsong is not complex enough to be a proper model species. This standpoint is mainly based on theoretical work, grounded in Chomsky's theory of Universal Grammar (Hauser, Chomsky & Fitch, 2002; Yang et al., 2017).

Here, I will outline this debate, in order to answer the question: *is birdsong a suitable model for human language?* I will not just describe the debate, but will also assess it from a linguistic perspective. In the first four chapters, I describe the debates and insights gained from it. Then, in the discussion, I critically review the debate from both a biological and a linguistic standpoint. This will lead to an integrated account of the similarities between birdsong and human language, encompassing work on semantics, phonology, syntax and vocal learning.

1 - Semantics

Semantics is the study of meaning associated with sounds, on the level of 'words'. In non-human species, this could be translated as 'the building blocks of phrases'. In humans, there are even smaller building blocks called phonemes (consonants and vowels), which make up words, which in turn make up phrases. It has been suggested that in birdsong, sounds that are much like phonemes directly make up the phrases (called motives), skipping the word level (Samuels, 2015). If an animal is capable of recognising a specific meaning when hearing a specific sound produced by another individual, and in turn, individuals are able to convey certain meanings through the production of certain sounds, there is communication going on (Shannon, 1948) and it can be said that the vocal communication of these individuals has a semantics. However, it is very hard to test whether this is actually the case.

The research on semantics in bird vocalisations started in the 1980s, when researchers became interested in alarm calls (Suzuki, 2016). The concept of alarm calls as such does not necessarily point at semantics, but the fact that birds can signal certain attributes of the predator through them (Templeton, Greene & Davis, 2005; Suzuki, 2014) does. It would be too quick to conclude that this means birds have at least some form of semantics, as it is not clear whether these alarm calls actually convey the meaning we often think they do. It might just be that these birds have paired a certain sound with a certain response, without the need for the complex cognitive mechanisms associated with semantics in humans. However, it might also be the case that the way alarm calls work are actually a very basic version of the system of semantics in humans. Another possibility is that birds merely use syntax (the way the building blocks of a phrase are organised) to convey meaning, without the individual sounds having specific meanings associated to them (Suzuki, 2016). This is going on in human language as well, on the level of phonetics: the individual phonemes are meaningless by themselves, but become meaningful when they are structured into words. Songbirds would then structure meaningless sounds into meaningful constructions by following rules observed in human syntax, but not rules observed in the way humans create words from phonemes.

Birdsong clearly has certain functions and elects certain behaviour (Catchpole & Slater, 2003). But do the sounds that make up a bird's song convey meaning? Is there a semantics of birdsong? This remains a complete mystery. In fact, almost all research on semantics in the vocal communication of birds has been focused on alarm calls, leaving birdsong out of the picture. Even if the building blocks of birdsong hold specific meanings, this has not been shown yet (Suzuki, 2016). This makes sense; proving some form of semantics is going on, is much easier when referential information is being communicated. Referential information refers to an object in the actual world. For example: bird A sees a predatory bird, and gives

an alarm call with the meaning “predatory bird”. If the meaning of this alarm calls indeed was “predatory bird”, bird B, which has not seen the predator, now not only knows that there is danger from the sky, but that the danger comes in the form of this predatory bird. Proving that this is the case remains hard, but is easier than proving there is semantics in birdsong, where it is much less clear what is being referred to, if anything is being referred to at all.

However, the research of semantics in birdsong is not completely hopeless. Computation analysis using methods from Artificial Intelligence might enable us to analyse birdsong in such a way that possible semantic information can be interfered. One such method has been suggested by Keisuke Daimon, Richard Hedley and Charles Taylor (2017). By using Dynamic Bayesian Networks, they identified relationships between context and song output in a certain species of songbird (*Vireo cassinii*). Furthermore, the great differences in patterns between individuals show that these birds might use highly local vocabularies in order to prevent eavesdropping by rivals. Such a method of analysis might prove interesting in further inquiry into the question whether the vocal communication of bird might code meaningful information, especially in complex sequences like birdsong.

2 - Phonology and prosody

Phonology is the study of the sounds that make up words in language. It thus concerns the meaningless components of the smallest level, that, in turn, form meaningful components. Prosody concerns itself with features of speech on the syllable level and above, like pitch, stress, intonation and rhythm. The neural pathways of these domains of human language are slowly becoming better understood. In this chapter, I will outline the ways bird phonology and prosody might be interesting for the understanding of human language.

The discussion of phonology in bird song tends to become obscured by a slight difference in meaning between bird phonology and human phonology. In humans, phonology and syntax are completely distinct areas, focusing on different levels. In the discussion of birdsong, syntactical structures are often discussed under the label of phonology (e.g. Yip, 2013; Samuels, 2015). This is not so surprising. Where human language has the levels of phonology (on which meaningless sounds become meaningful sound strings) and syntax (on which meaningful sound strings get arranged into a longer sound string with a more elaborate and specific meaning), birdsong is often thought to only have the phonological level. On this level, a difference between syllables (short note sequences) and song segments (sequences of syllables) is recognised, but the used terminology does not always reflect this. In human language, syntax-like structuring is going on on both these levels. However, on the phonological level different processes like assimilation of sounds, stuttering and prosodic effects are important as well, and these processes have been observed in several species of songbirds (Yip, 2013). Because of this, both the phonological and the syntactical are interesting in birdsong. However, the terms ‘syntax’ and ‘phonology’ are often mixed up in the discussion of syntactical features of birdsong.

2.1 - Similarities between bird and human phonology and prosody

Due to a combination of homologies of neural structures, combined with convergent evolution, phonology between birds and humans is very comparable on genetic, neuroanatomical and behavioural levels (Samuels, 2015). This becomes apparent, for example, in the hierarchical structure of both human language and birdsong (discussed in more detail in chapter 3) and the way surrounding sounds influence the way a sound is produced (Wohlgemuth et al., 2010), called “assimilation” in humans. However, there are differences between human and songbird phonology as well. Where human phonemes (vowels and consonants) are composed of certain features, like being voiced or unvoiced, the notes of birdsong seem to show a greater difference between them (Samuels, 2015). This has implications for the way these sounds must be parsed by the brain. Ultimately, human phonology and the phonology of songbirds seem to be very much alike on the

computational level, thus on the level of the underlying mechanisms, but less so on the representational level, which is the way this amounts to behaviour (Samuels, 2015).

2.2 - Phonological development

The perception and recognition of different sounds are at the very basis of language development in human infants. This process is shown to start even before birth and enables a child to recognise, and later use, the units of which its language is made up (Lust, 2006). It would make sense that, in order for birdsong development to be a good model for human language development, a similar kind of sensitivity to the basic units of a vocal communications system should be present, ideally with a similar process of acquisition as well.

It has been shown that in zebra finches (*Taeniopygia guttata*), segmentation using pauses enhances vocal learning in a way that is similar to the way human children use segmentation when learning language (Spierings, De Wege & Ten Cate, 2015). However, in this experiment adult zebra finches were used, even though zebra finches are shown to have a sensitive period for song-learning, similar to the critical period in humans (Gobes, Jennings & Maeda, 2017), which limits the time human children have for learning a language through mere exposure. As the neurological underpinnings of the phonological development of human children are relatively well understood (Lust, 2006), it would be interesting to see experimental work in songbirds on the phonological developments of their hatchlings. This way, the similarities and differences between bird and human phonology could become even more clear.

Despite the current lack of experimental evidence of songbirds going through a phonological development similar to human infants, the phonological system of songbirds remains interesting for researchers interested in human language. Zebra finches have been shown to be able to discriminate between human words that only differ on the vowel level (for example *taak* vs. *tak*), displaying the same cue weighting strategy as human test subjects (Ohms et al., 2012). Furthermore, the same research shows that zebra finches that had no experience with human speech, were especially sensitive to frequencies found within the range of their natural songs, suggesting that a sensitisation process, much like the way human infants become sensitive to the sounds of their mother tongue, might be present in songbirds as well. All in all, this suggests that the songbird phonological system might be highly similar to the human system, though more research is necessary to draw further conclusions.

2.3 - Prosody

In humans, prosody is an essential component of language, allowing speakers to call attention to certain parts of what they are saying, making the meaning of their words clearer to listeners. Because of this, prosody intertwines with many levels of human language, despite being produced on the phonological level. (Mol et al., 2017). However, certain prosodic effects have been observed in species of songbirds, including harmonic emphasis, tempo changes, rhythm and silent beats (Yip, 2013). In addition, at least some species of songbird appear to be sensitive to prosodic features of human speech and can discriminate between the same utterance with different prosodic effects (Naoi et al., 2012). Importantly, prosody seems to play an important role in the segmentation of songs as part of vocal learning in Bengalese finches (Takahasi, Yamada & Okanoya, 2010), just like it does in human language acquirement (Johnson & Seidl, 2009). Thus, despite possible different usage of prosody between humans and birds, the computational basis might highly alike.

This hypothesis is further supported by experimental evidence that songbirds and humans share a sensitivity to certain prosodic effects, including pitch, amplitude and duration. Moreover, zebra finches were shown to respond more strongly to prosodic pattern than to syntactic structure, whereas humans were shown to use one of both, based on personal preference (Spierings & Ten Cate, 2014).

Concluding, recent work in prosody in songbirds has highlighted that this is a potentially interesting feature of birdsong, but remains under-studied.

3 - Syntax

Syntax is the study of how (linguistic) units - such as words - are grouped into hierarchical structures following abstract rules. In human language, the term is used specifically to describe how words (linguistic units) form phrases (hierarchical structures) using grammar (abstract rules). In the discussion of birdsong, syntax concerns itself with the way notes are grouped into syllables and motives (Berwick et al., 2011). For birdsong to be a suitable model organism of human language, the syntactical structures it displays must be analogous to the syntactical structures found in human language. In this chapter, I will focus on the biggest discussion points regarding the investigation of syntax in birdsong: artificial grammar studies and syntactic complexity.

3.1 - Artificial grammar studies

Artificial grammar studies are often used to assess the ability of human and non-human animals to learn certain grammatical rules. In these experiments, subjects are exposed to an example set of string that has been generated using certain grammatical rules. Then, subjects are tested on their ability to discriminate between novel strings that are grammatical and ungrammatical, according to the tested syntactic rules (Beckers et al., 2017). These studies lie at the core of the inquiry into the syntactic abilities of animals, including songbirds. However, an enormous problem with this type of experiment is proving that the experiment indeed did test the syntactic ability of the test subject, and not some other cognitive ability (Ten Cate, 2014; Beckers et al., 2012). Another possible problem is that it is hard to distinguish between rules that were learnt because of grammatical rule learning and rules that were learnt because of acoustic biases, mainly increased acoustic similarity between the familiarisation stimuli and the grammatically correct testing stimuli. (Beckers et al., 2017).

However, that does not mean that this type of study is not fruitful at all. For example, artificial grammar protocols have shown that several species of mammals and birds have a categorical perception of sounds, meaning that they do not just recognise that exact sound, but also variants the a different within certain boundaries (Ten Cate, 2014).

Artificial grammar studies can also be used to compare the syntactical abilities of birds to those of humans. This does not necessarily negate the possible problems of this type of study – importantly uncertainty of the strategy the tested subjects used to come to their conclusions – but a great advantage is the ability to directly compare humans to non-humans. An example is an experiment in which both humans and zebra finches had to discriminate between grammatical and ungrammatical (according to the syntactical rules presented in the experiment) strings composed of zebra finch song elements. Zebra finches turned out to be able to discriminate between strings with a different order. This was not done by rote memorisation, but by generalisation, much like the way human infants do this.

However, this generalisation seems to be limited, as the birds were not able to transpose this generalisation to strings of the same form, but with a different surface level (Chen, Van Rossum & Ten Cate, 2014). The same was shown to be true for human infants in an earlier experiment (Gerken, 2006). The zebra finches were compared to adult humans, who were able to expand the learnt rules to strings with a different surface level, most likely because of their higher level of experience with rule abstraction (Chen, Van Rossum & Ten Cate, 2014).

3.2 - Syntactic complexity

At the heart of the discussion of birdsong syntax lies the distinction between *finite-state grammar* (FSG) and *context-free grammar* (CFG), as described in *Three models for the description of language* by Noam Chomsky (1956). According to Chomsky, different types of grammars are possible, each being able to generate phrases with a higher level of complexity. A finite-state grammar would be the lowest level, and therefore generate the most simple grammar, followed by context-sensitive, context-free and ultimately recursively enumerable grammar. Each next level expands on the levels below it (Figure 1), adding more rules. For example, the ability to embed a phrase into an existing phrase is associated with context-free grammar. In the discussion of syntactic complexity of birdsong, these terms are used to weigh the complexity of the syntax, and with that its comparability to the syntax of human language.

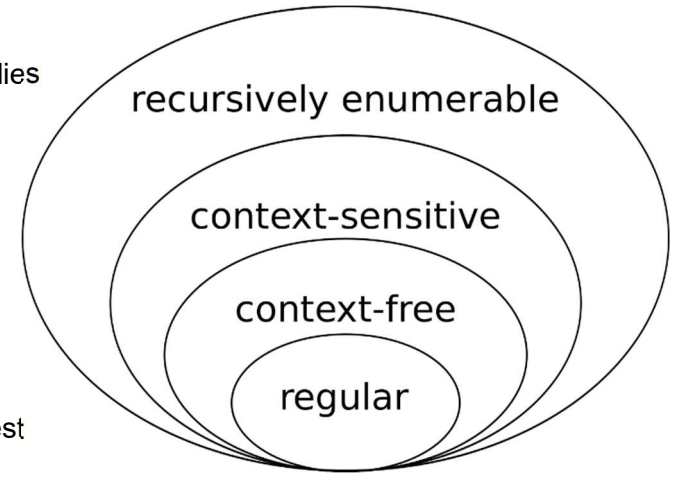

Figure 1 - A schematic representation of the Chomsky hierarchy (Chomsky, 1956). Source: wikicommons

When Kentaro Abe and Dai Watanabe published their results on syntactic rule discrimination by Bengalese finch (*Lonchura striata* var. *Domestica*) in 2011, this immediately sparked a vivid discussion. Up to that point, it was agreed upon that birdsong showed some level of syntactic complexity, but not whether this meant that songbirds possess the ability to structure auditory information through a computational system similar to that of humans. Human infants acquire the grammar of their mother tongue through exposure after their birth (Lust, 2006). In their publication, Abe & Watanabe provide evidence for Bengalese finches acquiring their syntactically complex songs postnatally as well. An important difference with the language development of humans is that these birds do not seem to have a critical period for this specific feature of their song, as the birds could even learn the songs for the first time as adults. Furthermore, their data suggest these birds have syntactic rules shared by their community, as they responded to certain subtle changes in the sequential ordering of syllables in their song, but not to others, and this was shared between individuals. This is further supported by neurological evidence: lesions in the anterior nidopallium (around

LMAN) led to the birds failing to recognised songs with manipulated syntactic structures, suggesting this brain area is involved in the discrimination of auditory information through analysis of syntactic functions. But most importantly: the grammar the researchers taught the birds was centre-embedding, embedding new phrases in the middle of an existing phrase and in that way increasing the complexity. This type of complex syntactic structures was long thought to only be achievable by humans.

The primary criticism on Abe & Watanabe comes from Beckers et al. (2012). Their main point of critique is that Abe & Watanabe claim that their experiment proves the birds used centre-embedding *syntax* to come to the right conclusion. However, it is possible that the birds did not use any grammar at all to do this, because the grammatically correct testing stimuli were similar enough to the familiarisation stimuli that is was possible to arrive at the right conclusion by memorising certain syllables and recognising whether these were present or absent during the test playback. The grammatically incorrect testing stimuli, on the other hand, showed a greater difference to the familiarisation stimuli. According to Beckers et al., (2012), this can all be done through cognitive abilities birds have been shown to possess. The fact that humans have been shown to have a lot of difficulty understanding centre-embedded phrases, only adds to the suspicion of the interpretation that Bengalese finches used grammar to recognise the right phrases in the test.

In conclusion, while it is conceivable that certain species of songbirds that sing syntactically complex songs might be in possession of such abilities, proving this is no easy feat. In tests of complex grammar, birds might use strategies to arrive at the right answer that do not rely on grammatical abilities. Because of this, it is hard to judge whether birds are capable of complex syntax in a way similar to humans. A better understanding of the neural underpinnings of syntax might help to prove that test subjects did use grammatical abilities to solve the test by blocking these pathways.

4 - Vocal learning

Vocal learning is among the prime reasons why birdsong has sparked the interest of researchers over evolutionary closer species like non-human primates (Petkov & Jarvis, 2012). Human language is highly complex, making learning it no easy feat. Yet, almost all humans acquire it without many problems, given enough input. This process consists of several subprocesses, some of which have been studied in quite a lot of detail. Furthermore, the examining of child language acquirement has led to a good part of our understanding of human language (Lust, 2006). Songbirds learning their repertoire do this through a complex process as well (Catchpole & Slater, 2008). In this chapter, I will further zoom in on two subprocesses associated with both human language acquirement and song learning in songbirds. These are: the critical period and learning biases. Another shared feature of the vocal learning of humans and songbirds, segmentation, was briefly discussed in section 2.2 (phonological development).

4.1 - The critical period

Lenneberg's critical period is a theory that states that human children can only learn their language during a certain period, starting at two years of age and lasting until the start of puberty (Lenneberg, 1967). It is hypothesised that during the critical period, certain areas of the brain involved with specific facets of language are more sensitive, allowing children to learn language through mere exposure. This seems to be true for phonology: newborns are still sensitive to all possible sounds in human language, but between six and twelve months, lose their sensitivity to sounds from languages they are not exposed to. For other facets of language, the theory is much harder to test, because depriving a child from language input when that could cause impaired language skills is rather unethical (Lust, 2006).

Songbirds learning to sing often display critical periods as well, though some species are *open-ended learners*, meaning they can acquire their song anytime in their life (Catchpole & Slater, 2008). Zebra finches are a well-studied example of *close-ended learners/age-limited learners*, who do display a sensitive period, possibly comparable to the critical period in humans. These birds do not start acquiring their song until the 25th day post-hatch, and learn it up to day 90 post-hatch (Immelmann, 1969). This learning process consists of two phases that do not overlap. During the first phase, the sensory phase, the bird listens and memorises the song from the tutor. Then, during the second sensorimotor phase, the bird learns to produce the song by actively comparing and gradually matching the memorised song of the tutor, as shown by deafening experiments (Doupe & Kuhl, 1999). Especially this last phase is interesting from a human language perspective, as the gradual matching of the song to the memorised template, seems to resemble babbling observed in human infants (Mol et al., 2017; Doupe & Kuhl, 1999).

4.2 - Learning biases

Another interesting feature of vocal learning is learning biases. Most, if not all, of biological work on human language reviewed here has been conducted with Chomsky's theory of Universal Grammar in mind. This theory states that there is an internal language faculty that know the rules of human grammar, making it easier for children to learn the structure of their native language (Chomsky, 1986). In other words, the theory states that certain grammatical rules are inborn. However, Universal Grammar remains far from uncontroversial. While many linguists these days will not deny that there is some biological basis to language, a good amount of them does not believe Chomsky's theory is a satisfying account of what this biological basis is like. One of the alternative hypothesises argues that certain structures in human language are not structures by hard-wired, genetically imprinted rules, but instead by more general learning and processing biases of the mind and motor-sensory systems (Christiansen & Chater, 2008).

Research on learning biases in songbirds could help answer the question of what theory seems more likely. One such study has shown that zebra finches that have been tutored using randomised tutor-sequences develop normal and non-random song. The developed songs were highly similar to zebra finch songs encountered in the wild, suggesting the convergence was a result of learning biases. Interestingly, these patterns are also observed in human language and music. Possibly, this is the result of the process governing this pattern forming being highly conserved among vertebrates (James & Sakata, 2017). This would imply that at least some part of the zebra finch song system could resemble a basic version of the human language system, at least on the level of syntax. This would make the zebra finch an interesting model for human language.

5 - Discussion

Birdsong is an interesting phenomenon, especially because of its similarity to human language when it comes to simple syntax, phonological categories and the patterns and pathways of vocal learning (Jarvis, 2009). The apparent lack of semantics associated with birdsong might make it not seem like a good model for human language. However, this apparent simpleness of the systems of birdsong might be just what is needed to further the understanding of language-like systems. Human language is highly complex, with many variables that are almost impossible to isolate. A simpler model system might enable linguists to get to the basis of our human language system.

But can birdsong be such a system? This is a more complex question than it might seem, as both birdsong and human language are made-up out of many different processes, that can be categorised under the domains semantics, phonology and syntax, with vocal learning governing how these domains are acquired. To complicate matters even further, there are great interspecies differences in the learning and realisation of birdsong, for example the great syntactic complexity of Bengalese finch song, the enormous repertoire of nightingales and the European starling's ability to acquire their song at any moment in their life, as opposed to the sensitive period for song acquirement as observed in the zebra finch. In order to assess the suitability of birdsong as a model for human language, all these things must be taken into consideration. Because of this, I will discuss each of these processes separately, including suggestions for further research.

5.1 - General discussion

The discussion on the value of birdsong as a model for the study of human language is often held from a so-called "biolinguistic approach" (Berwick and Chomsky, 2011; Yang et al., 2017). This name is misleading, as it might be easily confused with the field of biolinguistics, in which language is studied with a focus on its underlying biology, but not necessarily with the theory of Universal Grammar as a basis (e.g. Christiansen & Chater, 2008). The "biolinguistic approach" as described by Berwick and Chomsky (2011), however, does take Universal Grammar as a starting point (see Boeckx & Grohmann, 2007 as well). Because of this, the biological study of language, especially with regard to possible animal models, is shaped to a high degree by the theory of Universal Grammar, which is far from undisputed within its field of origin, linguistics (e.g. Ibbotson & Tomasello, 2016; Lin, 2017). Thus, much of this research is based on the idea that at the basis of the language system, there is a rich set of grammatical rules, that help parse the input. This is a very specific version of how language could have biological foundations; other explanations are very much possible as well, but largely ignored because of the ubiquitousness of Universal Grammar.

This poses a big problem in this kind of research, and thus in assessing the suitability of birdsong as a model for human language. In many of the review articles, birdsong is deemed too simple to be a suitable model, because it does not resonate with the theory of Universal Grammar enough (e.g. Berwick et al., 2011; Bolhuis et al., 2018). This may prove to be a loss on two sides: (1) an interesting line of research might not be followed, because on the basis of a heavily discussed claim it is not deemed interesting enough, and (2) the discussed claim cannot be assessed using these observations, because the interpretations following from them are deemed dubious by the discussed claim. Thus, the conviction surrounding Universal Grammar in the discussion of language-like systems in non-human animals, might not be helpful in determining whether these animals are interesting models to deepen the understanding of human language, but might instead be holding the whole field back.

One more general remark: as noted above, the systems of birdsong show great differences between species of songbird. Therefore, it does not make sense to decide on a single species to use as a model organism. Rather, it should be assessed the song of which species seems to resemble a certain aspect of human language the best, for example zebra finches for their sensitive period, and Bengalese finches for the complex syntactic structure of their songs. Studying the same aspect in several species and comparing the results to each other and humans might be an interesting approach as well (Brenowitz & Beecher, 2005).

Altogether, birdsong does seem to pose an interesting model for the understanding of human speech. Certain features of birdsong appear to closely resemble human language (in what ways is explained in the following paragraphs). Furthermore, there are practical advantages: songbirds are relatively easy to keep in the lab, much easier than other species that might be interesting models as well, like whales (Janik, 2014) and bats (Vernes, 2017). Because of this, the workings of their song systems are relatively well understood, though more research is necessary to determine how these exactly relate to the human language system. However, it should first be carefully studied which species would be the best model for which aspects of these systems. Also, the strong focus on Universal Grammar as an understanding of human language at the biological level might be reconsidered, allowing for alternative explanations.

5.2 - Semantics

At this point, birdsong is not a suitable model for human semantics. Because of the difficulties surrounding the study of semantics in non-human animals – in particular, determining whether meaning has been communicated, and if so, what meaning exactly – it is highly unlikely that an animal model would be suitable in this domain at all.

However, more research on the question whether birds have a semantics is more than welcome. There is quite some discussion going on about the status of syntax without semantics (e.g. Yip et al., 2013; Mol et al., 2017), and a better understanding of whether birds have a form of semantics might help guide this discussion, and further our understanding of the ways in which the avian system of vocal communication is comparable to that of humans.

5.3 - Phonology

Phonology is a domain in which birdsong could be an interesting modal for human language. Birds have been shown to use certain phonological and prosodic features to structure auditory strings (Mol et al., 2017; Spierings and Ten Cate, 2014; Ohms et al., 2011). Moreover, the fact that songbirds are shown to have phonetic categories just like humans (Ten Cate, 2014), means that they could be interesting in that regard as well.

Yet, here too additional research is necessary to discover how songbirds can be interesting to the study of human language. It is very interesting that songbirds have phonological categories, but it remains unknown how these are acquired. As this is relatively well-understood in humans (Lust, 2006), this could be an interesting area for some direct comparisons between the song system of songbirds and the language system of humans. Prosody might prove to be another area for fruitful comparisons. Birds and humans seem to share sensitivities in this regard (Spierings & Ten Cate, 2014), but this remains a relatively underdeveloped area of research.

5.4 - Syntax

The study of syntax is a particularly complicated undertaking. Artificial grammar learning (AGL) studies seem to be the only way to assess whether the syntactic system of songbirds is sensitive to abstract rules in a way comparable to human syntax, yet it remains difficult to prove the birds arrived at the right answer because of grammatical abilities, instead of other cognitive abilities (Beckers et al., 2012). This makes studying the limits of the syntactic system of songbirds difficult, and because of this, it is hard to determine whether those systems are comparable to the syntactic systems of humans. Yet, this does not mean that this type of study is futile. However, AGL protocols should be carefully designed, trying to include testing strings that could reveal non-grammatical solving strategies. In addition, a better understanding of the neural underpinnings of syntax might help to prove that test subjects did use grammatical abilities to solve the test by blocking these pathways.

Because Universal Grammar is mainly a theory on syntax, this domain is especially harmed by the focus on Universal Grammar explanations of language-like systems. It might prove to be fruitful to reanalyse the song of several species of songbird with different theories

about the biological origin of language in mind (e.g. Christiansen & Chater, 2008). Importantly, even when holding on to a Universal Grammar view on language, the goal should not be to find songbirds with syntactic abilities that seem to be on par with the syntactic complexity observed in human language. Rather, a simpler syntax might still prove to be an interesting model, as these simple rules might lie at the basis of more complex rules found in human syntax. An interspecies-analysis might be interesting too, as understanding the syntactical differences between the song of several species of songbird might tell us more about how songbird syntax could relate to human syntax, or perhaps even about the syntactic differences between human languages.

Finally, the future developments in this domain might prove to be closely related to developments in the domain of semantics. In humans, syntax is essential in the constructing of the meaning of utterances. Whether this is the case in songbirds as well, remains a complete mystery. It could be that there is no meaning, and the observed syntax is simply a result of the structure of the song system. Alternatively, it could be that syntax in birdsong is actually important in the perceived meaning of a song fragment, which would have great implications on the way we understand avian syntax. This could for example mean that avian syntax, like human syntax, is compositional (Suzuki, Wheatcroft & Griesser, 2016; Suzuki, Wheatcroft & Griesser, 2018; Bolhuis et al., 2018).

5.5 - Vocal learning

Vocal learning in songbirds has been the object of studies for many decades. This research has shown that some, but not all, songbirds have a sensitive period in which they are able to learn their song (e.g. Immelmann, 1969). This sensitive period might be comparable to the hypothesised critical period in humans. An interesting way forward might be more direct comparisons between the biology underlying this sensitive period and the supposed critical period in humans. However, because it appears that the different domains of language each have different critical periods in humans (Lust, 2006), the best course of action might be to study the underlying mechanisms of possible sensitive periods in connection with the general investigation into the relation between these domains in the songbird song system and human language system.

Furthermore, in recent years, studies into the vocal learning of songbirds have become more focused on the neuroanatomical and genetic similarities in the vocal learning of songbirds and humans (e.g. Petkov & Jarvis, 2012; Austin et al., 2012; Pfenning et al., 2014). More research directly comparing songbirds and humans in this way might be very interesting, especially if they follow up on work done at the behavioural level of the different domains. This might not only help us understand the underlying biology of both systems better, but also the evolution of these systems.

A study showing learning biases for certain syntactic structures in zebra finches (James & Sakata, 2017) opens up an interesting new terrain for research. Such learning biases might help us understand the biological foundations of language better, and might even help to resolve the debate surrounding Universal Grammar. Further research into this could change the direction of all of biolinguistics.

References

Abe, K. & Watanabe, D. (2011) Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nature Neuroscience 14*(8), 1067-1074. doi:10.1038/nn.2869

Austin, A.T., Miller, J.E., Fraley, E.R., Horvath, S. & White, S.A. (2012). "Molecular Microcircuitry Underlies Functional Specification in a Basal Ganglia Circuit Dedicated to Vocal Learning." *Neuron* 73, no. 3: 537-552

Berwick, R., Okanoya, K., Beckers, G., & Bolhuis, J. (2011). Songs to syntax: The linguistics of birdsong. *Trends in Cognitive Sciences, 15*(3), 113-21. doi:10.1016/j.tics.2011.01.002

Berwick, R. C., & Chomsky, N. (2011). The biolinguistic program: The current state of its development. In A. Di Sciullo & C. Boeckx (Eds.), *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty* (pp. 19-41). Oxford: Oxford University Press.

Beckers, G., Bolhuis, J., Okanoya, K., & Berwick, R. (2012). Birdsong neurolinguistics: Songbird context-free grammar claim is premature. *Neuroreport, 23*(3), 139-45. doi:10.1097/WNR.0b013e32834f1765

Beckers, G., Berwick, R., Okanoya, K., & Bolhuis, J. (2017). What do animals learn in artificial grammar studies? N*euroscience and Biobehavioral Reviews, 81*(Part B), 238-246. doi:10.1016/j.neubiorev.2016.12.021

Brenowitz, E., & Beecher, M. (2005). Song learning in birds: Diversity and plasticity, opportunities and challenges. *Trends in Neurosciences, 28*(3), 127-32. doi:10.1016/j.tins.2005.01.004

Boeckx, C., & Grohmann, K. K. (2007). The biolinguistics manifesto. *Biolinguistics, 1*, 001-008.

Bolhuis, J. J., Beckers, G. J., Huybregts, M. A., Berwick, R. C., & Everaert, M. B. (2018). Meaningful syntactic structure in songbird vocalizations?. *PLoS biology, 16*(6), e2005157. doi:10.1371/journal.pbio.2005157

Catchpole, C. K., & Slater, P. J. B. (2008). *Bird song: Biological themes and variations* (2nd ed.). New York, NY, US: Cambridge University Press.

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*, 157-64. doi:10.1016/j.conb.2014.07.019

Chen, J., Van Rossum, D., & Ten Cate, C. (2014). Artificial grammar learning in zebra finches and human adults: XYX versus XXY. *Animal Cognition, 18*(1), 151-164. doi:10.1007/s10071-014-0786-4

Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on information theory, 2*(3), 113-124.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. New York, N.Y.: Praeger.

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and brain sciences, 31*(5), 489-509.

Condro, M. C., & White, S. A. (2014). Recent Advances in the Genetics of Vocal Learning. *Comparative Cognition & Behavior Reviews, 9*, 75-98.

Daimon, K., Hedley, R., & Taylor, C. (2017). Semantic Inference of Bird Songs Using Dynamic Bayesian Networks. In *AAAI Conference on Artificial Intelligence*. Retrieved from https://aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14256

Darwin, C. (2016). *The descent of man, and selection in relation to sex.* Dinslaken: Anboco. (Original work published 1871).

Doupe, A.J., & Kuhl, P.K. (1999). Birdsong and human speech: common themes and mechanisms. *Annual review of neuroscience*, *22*(1), 567-631.

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

Gobes, S., Jennings, R., & Maeda, R. (2017). The sensitive period for auditory-vocal learning in the zebra finch: Consequences of limited-model availability and multiple-tutor paradigms on song imitation. *Behavioural Processes, 2017 Jul 23* (online only). doi:10.1016/j.beproc.2017.07.007

Hauser, M.D., Chomsky, N.W. & Fitch, 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

Ibbotson, P., & Tomasello, M. (2016). What's Universal Grammar? Evidence rebuts Chomsky's theory of language learning. Scientific American. Retrieved on February 7 2019, from http://cogsys.sites.olt.ubc.ca/files/2016/09/Whats-universal-grammar-Eviden-ce-rebuts-Chomsky%E2%80%99s-theory-of-la.pdf

Immelmann, K. (1969). Song development in the zebra finch and other estrildid finches. *Bird vocalizations,* 61-77.

James, L., & Sakata, J. (2017). Learning biases underlie "universals" in avian vocal sequencing. *Current Biology : Cb, 27*(23), 3676-3682. doi:10.1016/j.cub.2017.10.019

Janik, V. M. (2014). Cetacean vocal learning and communication. Current Opinion in Neurobiology, 28, 60-65.

Jarvis E. D. (2007). Neural systems for vocal learning in birds and humans: a synopsis. *Journal of ornithology, 148*(1), 35-44.

Johnson, E. K., & Seidl, A. H. (2009). At 11 months, prosody still outranks statistics. *Developmental science, 12*(1), 131-141.

Lenneberg, E. H. (1967). The biological foundations of language. *Hospital Practice, 2*(12), 59-67.

Lin, F. Y. (2017). A refutation of Universal Grammar. *Lingua, 193*, 1-22.

Lust, B. (2006). *Child language : Acquisition and growth* (Cambridge textbooks in linguistics). Cambridge: Cambridge University Press. (2006). doi:10.1017/CBO9780511803413

Mol, C., Chen, A., Kager, R. W., & ter Haar, S. M. (2017). Prosody in birdsong: A review and perspective. *Neuroscience & Biobehavioral Reviews, 81*, 167-180. doi:10.1016/j.neubiorev.2017.02.016

Naoi, N., Watanabe, S., Maekawa, K., & Hibiya, J. (2012). Prosody discrimination by songbirds (Padda oryzivora). *PLoS One, 7*(10), e47446. doi:10.1371/journal.pone.0047446

Ohms, V. R., Escudero, P., Lammers, K., & ten Cate, C. (2011). Zebra finches and Dutch adults exhibit the same cue weighting bias in vowel perception. Animal cognition, 15(2), 155-61. doi:10.1007/s10071-011-0441-2

Petkov, C. I., & Jarvis, E. D. (2012). Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. *Frontiers in evolutionary neuroscience, 4(*12). doi:10.3389/fnevo.2012.00012

Pfenning, A., Hara, E., Whitney, O., Rivas, M., Wang, R., Roulhac, P., . . . Jarvis, E. (2014). Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science, 346*(6215), 1256846-1256846. doi:10.1126/science.1256846

Samuels B.D. (2015). Can a bird brain do phonology?. *Frontiers in psychology, 6,* 1082. doi:10.3389/fpsyg.2015.01082.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell system technical journal, 27*(3), 379-423.

Spierings, M. J., & ten Cate, C. (2014). Zebra finches are sensitive to prosodic features of human speech. *Proceedings. Biological sciences, 281*(1787), 20140480. doi:10.1098/rspb.2014.0480

Spierings, M., de Weger, A., & Ten Cate, C. (2015). Pauses enhance chunk recognition in song element strings by zebra finches. *Animal cognition, 18*(4), 867-74. doi:10.1007/s10071-015-0855-3

Suzuki, T. N. (2014). Communication about predator type by a bird using discrete, graded and combinatorial variation in alarm calls. *Animal Behaviour, 87*, 59-65.

Suzuki, T.N. (2016) Semantic communication in birds: evidence from field research over the past two decades. *Ecological Research* 31, 3: 307-319. doi:10.1007/s11284-016-1339-x

Suzuki, T. N., Wheatcroft, D., & Griesser, M. (2016). Experimental evidence for compositional syntax in bird calls. *Nature communications, 7*, 10986. doi:10.1038/ncomms10986

Suzuki, T. N., Wheatcroft, D., & Griesser, M. (2018). Call combinations in birds and the evolution of compositional syntax. *PLoS biology, 16*(8), e2006532. doi:10.1371/journal.pbio.2006532

Takahasi, M., Yamada, H., & Okanoya, K. (2010). Statistical and prosodic cues for song segmentation learning by Bengalese finches (*Lonchura striata* var. *domestica*). *Ethology, 116*(6), 481-489.

Tchernichovski, O., & Marcus, G. (2014). Vocal learning beyond imitation: Mechanisms of adaptive vocal development in songbirds and human infants. *Current Opinion in Neurobiology, 28*, 42-47. doi:10.1016/j.conb.2014.06.002

Templeton, C. N., Greene, E., & Davis, K. (2005). Allometry of alarm calls: black-capped chickadees encode information about predator size. *Science, 308*(5730), 1934-1937.

Vernes S. C. (2016). What bats have to say about speech and language. *Psychonomic bulletin & review, 24*(1), 111-117.

Yang, C., Crain, S., Berwick, R., Chomsky, N., & Bolhuis, J. (2017). The growth of language: Universal Grammar, experience, and principles of computation. *Neuroscience and Biobehavioral Reviews, 81*(Part B), 103-119. doi:10.1016/j.neubiorev.2016.12.023

Yip, M. (2013). Structure in Human Phonology and in Birdsong: A Phonologist's Perspective. In Bolhuis, J.J., & Everaert, M. (Eds.) *Birdsong, Speech and Language: Converging Mechanisms* (pp. 181-208). Cambridge, Massachusetts: The MIT Press.

Wohlgemuth, M. J., Sober, S. J., & Brainard, M. S. (2010). Linked control of syllable sequence and phonology in birdsong. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 30*(39), 12936-49.