The structure of an avian syllable syntax network

Viviane Deslandes [a,b,*], Luiz R.R. Faria [a,c], Marcelo E. Borges [a], Marcio R. Pie [a,b]

[a] *Laboratório de Dinâmica Evolutiva e Sistemas Complexos, Departamento de Zoologia, Universidade Federal do Paraná, Curitiba, PR, Brazil*
[b] *Pós Graduação em Ecologia e Conservação, Universidade Federal do Paraná, Curitiba, PR, Brazil*
[c] *Instituto Latino-Americano de Ciências da Vida e da Natureza, Universidade Federal da Integração Latino-Americana (UNILA), Brazil*

ARTICLE INFO

ABSTRACT

A common result in recent linguistic studies on word association networks is that their topology can often be described by Zipf's law, in which most words have few associations, whereas a few words are highly connected. However, little is known about syntactic networks in more rudimentary communication systems, which could represent a window into the early stages of language evolution. In this study, we investigate the syntactic network formed by syllable associations in the song of the oscine bird *Troglodytes musculus*. We use methods recently developed in the context of the study of complex networks to assess topological characteristics in the syntactic networks of *T. musculus*. We found statistically significant evidence for nestedness in the syllable association network of *T. musculus*, indicating network organization around a core of commonly used notes, small-world features, and a non-random degree distribution. Our analyses suggest the possibility of a balance between the maintenance of core notes and the acquisition/loss of rare notes through both cultural drift and improvisation. These results underscore the usefulness of investigating communication networks of other animal species in uncovering the initial steps in the evolution of complex syntax networks.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

The study of complex networks has been increasingly used in linguistic research to describe the syntactic structure of languages (e.g. Cavalli-Sforza et al., 1988; Solé et al., 2010). In this approach, words are described as vertices in a graph, and their association in speech or writing is represented by an edge (Solé, 2005; Steyvers and Tenenbaum, 2005). Their topology is a function of the connectivity among its edges, leading to a variety of patterns, including random, small-world, and scale-free networks (Watts and Strogatz, 1998). In particular, word networks are often described as scale-free, with considerable variation in connectivity between their edges (Ferrer i Cancho and Solé, 2001; Albert and Barabási, 2002) such that most words are poorly connected and a few words account for a disproportionately large number of connections (Solé, 2005). Moreover, scale-free word association networks seem to emerge at a fairly discrete stage during child development from more elementary random networks

characteristic of toddler speech (Corominas-Murtra et al., 2009, but see Lipkind et al., 2013).

The structure of human syntactic networks is invariably complex, even in fairly simple societies (Cavalli-Sforza et al., 1988; Solé, 2005; Solé et al., 2010). As a consequence, it is difficult to envision the early stages of language evolution. An alternative approach to this issue would be to investigate communication systems of animals with more rudimentary syntactic networks. The physiological similarities in speech acquisition in humans and song development in oscine birds could suggest that avian song syllables networks might represent a valuable model system to investigate syntactic evolution in non-human animals (Doupe and Kuhl, 1999; Bolhuis et al., 2010). In addition to the ability to learn songs, vocalizations in oscine birds are complex and flexible, even allowing for the existence of local dialects (Podos and Warren, 2007; Catchpole and Slater, 2008), and for the adjustment of their syntax depending on the social context (Peters et al., 1980; Searcy et al., 1981; Kreutzer and Brémond, 1986; Balaban, 1988; Holland et al., 2000; Abe and Watanabe, 2011). For instance, there is variation in phrase arrangement between different populations of *Alauda arvensis* (Briefer et al., 2008), leading to a decrease in syllable sharing with geographical distance (Briefer et al., 2010). Despite the clear parallels between word associations in human language and syllable associations in bird songs, only two studies to date used network metrics
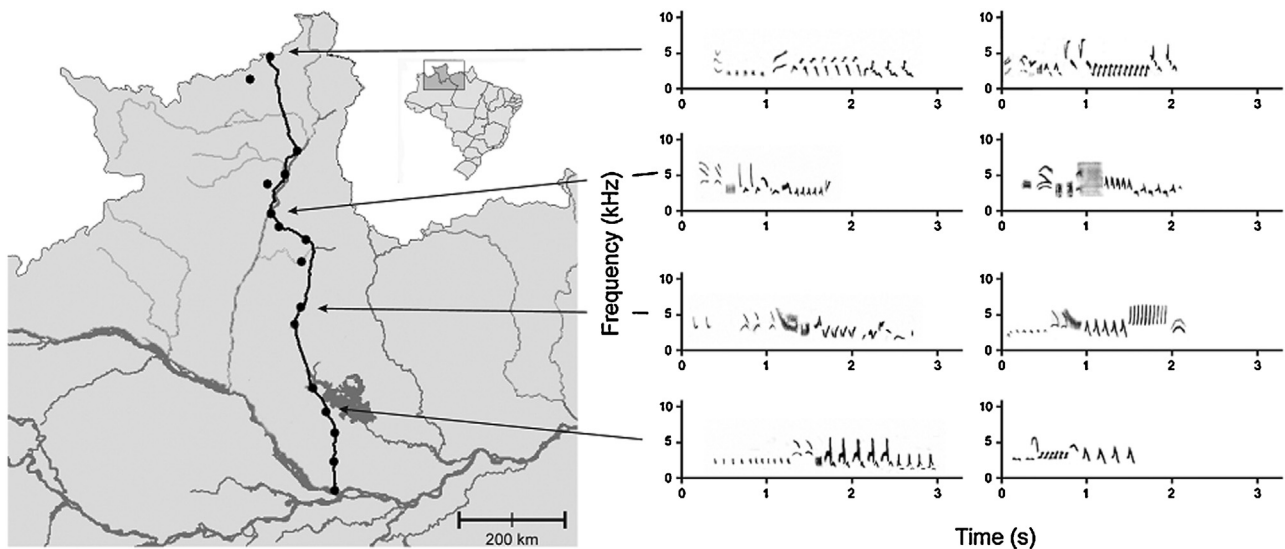
Fig. 1. Sampling locations and representative spectrograms based on recordings of *Troglodytes musculus*. Songs were recorded in 16 locations along an 1140-km transect following the BR-174 highway. Each spectrogram in the figure represents the song of a different individual.

to analyze the birdsong syntax (Briefer et al., 2010; Sasahara et al., 2012).

Troglodytes musculus is a sedentary neotropical oscine with a wide geographical distribution from southwestern Mexico to the Tierra del Fuego, living in open and semi-open habitats, forest clearings, and areas affected by human disturbance (Sick, 1997). Their song is composed of series high- and low-pitched syllables, with the entire sequence lasting between 2 and 9 s (Brewer, 2001; Kroodsma, 2005). Their syllable repertoire (number of different syllable types sung by this species) is vast and complex (Sick, 1997), and preliminary analyses indicated considerable variation in syllable succession, even in a single phrase, suggesting that this species is an ideal model to investigate the syntactic structure of syllable networks. In this study we tested if syllables networks in *T. musculus* depart from a simple random association of notes in their repertoire, as indicated by (i) significant nestedness, as expected for network organization around a core of specific syllables, and (ii) a frequency distribution of syllable use that deviates significantly from the expected in the case of random syllable associations.

2. Material and methods

2.1. Data collection

Songs were recorded from April to May and from September to December, 2006, in 16 locations along an 1140-km transect following the BR-174 highway, from Manaus-AM to Pacaraima-RR, Brazil, from 5:30 a.m. to 10:30 a.m. and 3:30 p.m. to 5:30 p.m. (Fig. 1). The average distance between sampling locations was 76 km, ranging from 30 to 260 km. Between 2 and 15 birds were recorded in each location, following a straight perpendicular transect in relation to BR-174, with a minimum interval of 200 m between each successive recording. No specific permissions were required to record the birds in these locations and neither manipulation nor experiments were conducted with the birds. We separated recordings by a minimum distance of 200 m to prevent recording the same bird twice, given that it is larger the expected distance between neighboring territories (Kendeigh, 1941; Tubaro, 1990). Songs were captured without playback, using a Sony TCM 5000 EV recorder equipped with a Sennheiser ME66 directional microphone. Each recording was digitalized using Edirol Audiocapture UA-5 and ADOBE AUDITION 1.0 (Adobe, 1992–2003), with resolution of 16 bits and sample

rate of 44.1 kHz. All recordings were deposited in the Bird Collection of the Instituto Nacional de Pesquisas da Amazônia (INPA).

2.2. Data analysis

All songs types recorded from each individual were analyzed. A song was defined as a sequence of syllables arranged in short time interval, separated by periods of silence, which was then followed by another song. Syllable types classified in spectrograms are exemplified in Fig. 2. All spectrograms were generated using Signal 4.0 (http://www.endes.com), with the following specifications: FFT: 256; Window: Hann and sampling frequency: 22.050 kHz. The recording and characterization of each syllable was made by the same researcher (V.D.). Although this type of classification can introduce subjectivity in our analysis, previous studies show that results using this method do not differ significantly of more advanced mathematical methods (Nowicki and Nelson, 1990; Podos et al., 1992; Botero et al., 2008). A numerical classification was assigned for each distinct syllable (ranging from "N1" to "N42") that was used in the construction of a matrix representing the sequence of syllables of each different song. Fig. 3 represents all the 42 syllable types found in this study. We confirmed the reliability of our classification system by randomly selecting 30 songs in our sample, which were then re-classified by the same researcher (VD) and later compared with the original classification. We found that over 88% of correspondence between both classifications (data not shown). In the next step, we used the matrix with syllables classification to generate two different networks: $MSeq_{42 \times 42}$ and $MQuant_{207 \times 42}$.

In graph theory, networks are depicted as an adjacency matrix, where lines i and columns j correspond to vertices or nodes, and elements a_{ij} correspond to the edges or links between two nodes. A value a_{ij} that is equal or greater than 1 indicates the presence of a link between two nodes, and a value equal to 0 indicates its absence. Edges can be directed if the network implies an order among the nodes, or undirected if it represents only an association between the elements. The network $MSeq_{42 \times 42}$ was built as a square matrix in which edges are weighted by the total frequency of links between two successive pair of syllable types, and directed according to their respective sequence. The preceding syllables were shown as columns and the subsequent syllables represented as rows. In the second network, named $MQuant_{207 \times 42}$, each line represents a
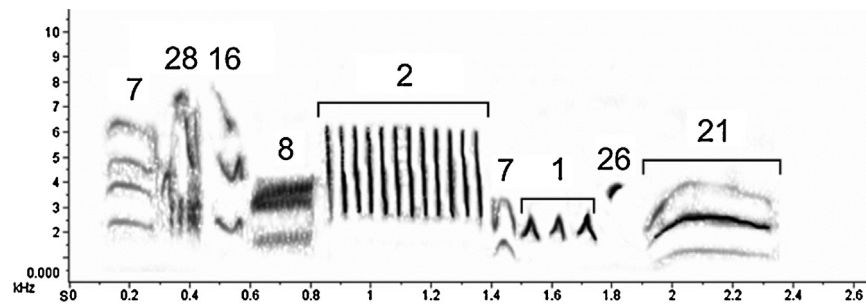
Fig. 2. Example of a spectrogram of *T. musculus.* The syllables are indicated by numbers. Different syllable types received different numbers and a sequence with the same number was named a phrase. This song shows nine syllable types, in which two types (classified as one and two) are forming phrases.

different song and each column corresponds to a syllable type. The edges were weighted according to the frequency of syllable use in each song.

We characterized the structure of *MSeq* according to four commonly used network metrics (Steyvers and Tenenbaum, 2005): average distance (L), diameter (D), clustering coefficient (C), and the degree distribution $P(k)$. L and D are closely related metrics

and are related to the concept of network path or geodesic distance (Bouttier et al., 2003). A path in a network is defined by the number of steps necessary to connect two distinct nodes. For example, a linear graph with four sequentially connected nodes requires three steps to reach the farthest nodes. L refers to the average of the shortest path lengths between all pairs of nodes in a network, while D refers to the maximum of these distances over all pairs of
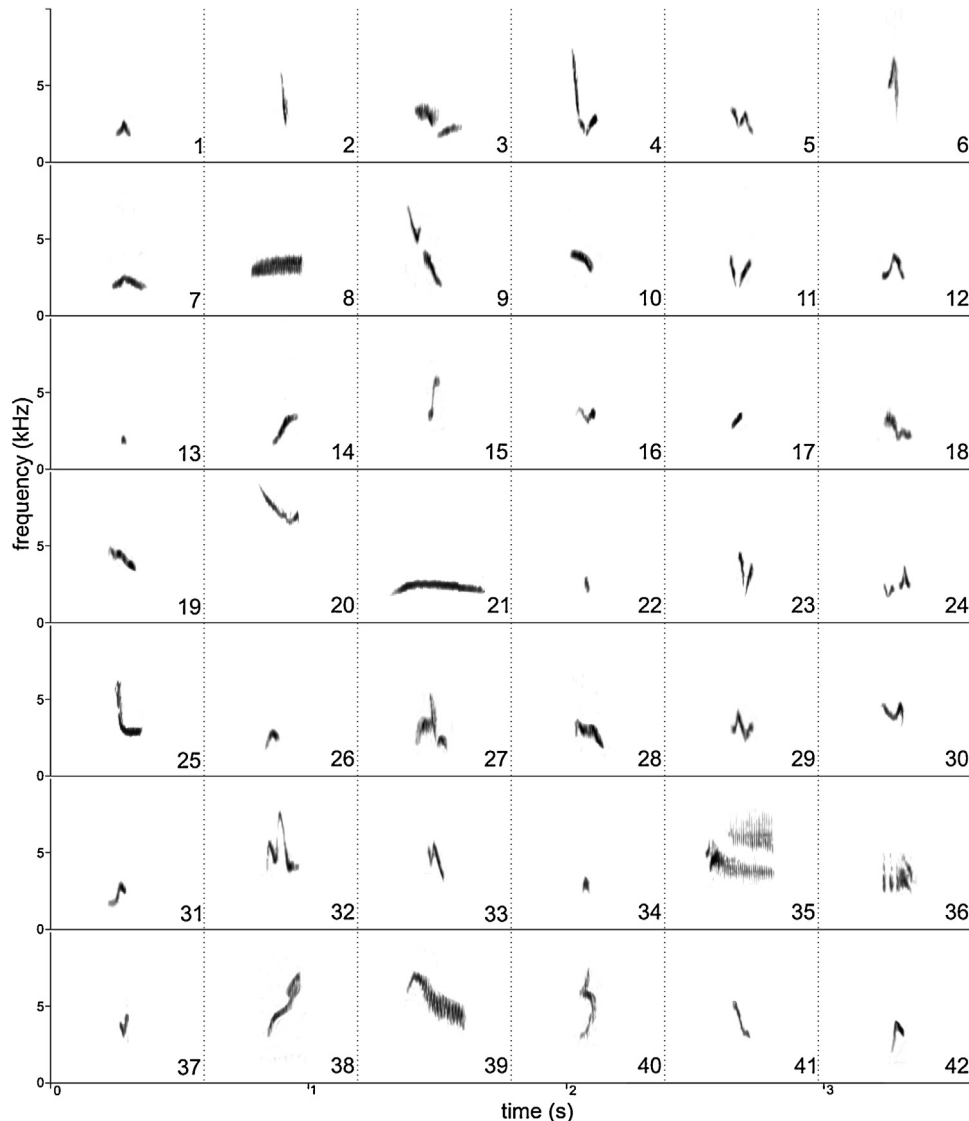

Fig. 3. Syllables types found in *T. musculus* songs sampled in this study. The original spectrograms used to select the syllables were generated with the same time and frequency resolution; therefore the syllable types shown in this figure keep their original format and size.

nodes in the graph. C is a measure of the proportion of neighboring nodes that can be reached through the nodes of other neighbors, calculated as the proportion of a focal node's neighbors who are themselves neighbors (Proulx et al., 2005). The clustering coefficient of a network is calculated by the arithmetical mean of the individual C_i for all vertices. Finally, $P(k)$ represents the probability that a randomly chosen node will have degree – the number of nodes that connect the focal node to other nodes – k (i.e., will have k neighbors) (see Steyvers and Tenenbaum, 2005; Proulx et al., 2005).

To test if the structure of the networks departed from a random topology, we assessed the extent to which L, D, and C were significantly different from the expected values of equivalent random networks. We compared the observed metrics with the mean value of 1000 simulated datasets using the packages IGRAPH(Csardi and Nepusz, 2006) and SNA (Butts, 2010) in R software. Random networks were generated according to the Erdös-Rényi model (Erdös and Rényi, 1959). Parameters were set equal to the number of vertices and edges of the original networks. The model creates an initial graph with N nodes and zero edges. At every step, a new edge is randomly assigned between two nodes. The process is repeated until it is reached a maximum number of edges M. The null expectations for L, D, and C were obtained using the resulting 95% confidence intervals. We calculated the index of small-world-ness (S^{ws}), according to Humphries and Gurney (2008). A network is considered to be small-world if $S^{ws} > 1$. $P(k)$ was plotted and compared against the degree distributions of all null models plotted simultaneously. This approach enables us to visualize whether the degree distribution overlaps a most probable range of expected distributions generated stochastically. We also assessed the fit of the degree distribution of this network by comparing it with three alternative models (exponential, power-law, and truncated power law) using the Akaike Information Criterion-AIC (Amaral et al., 2000). In order to examine the sequential structure of the song, we used a correlation analysis of the number of interactions of each syllable with other syllables, as well as with itself. We also used the frequency of syllable use across the *MQuant* to test for the existence of Zipf's law – a power-law distribution of syllable use – by comparing it to two alternative models (exponential and truncated power-law) using the Akaike Information Criterion-AIC (Amaral et al., 2000). According to Zipf's law, if elements are ranked according to their probability of occurrence, from commonest to the rarest, there must be an inverse monotonic relationship between rank of use and probability of occurrence (Briefer et al., 2010).

Given that both *MQuant* and *MSeq* are quantitative matrices, we used the WINE method implemented in the BIPARTITE R package (Dormann et al., 2008, 2009). WINE takes into account the weight of the interactions according to their relative frequencies. The nestedness index is calculated through the average distance from each matrix cell containing a link to the cell with the lowest marginal totals, in the packed matrix, using a weighted Manhattan distance. This metric was chosen mainly by their insensitivity to features such as matrix shape and size (Galeano et al., 2009). This method provides an important advance in relation to previous studies based only on incidence matrices and therefore is more powerful in uncovering the level of nestedness in a given matrix. Graph visualization was conducted using the packages SNA (Butts, 2010) and BIPARTITE in R (Dormann et al., 2008, 2009).

3. Results

The song repertoire of *T. musculus* was remarkably diverse, allowing for the identification of 42 distinct syllable types. Song variation within a location was comparable to that among different locations (Fig. 1). The number of syllable types recorded in any given song varied between 3 and 13 ($\bar{x} = 7.2$, $SD = 2.1$, $N = 207$ songs),

Table 1
Relative fit of the *MQuant* to Zipf's according to three models: Exp: Exponential; PL: Power Law and TPL: Truncated Power Law.

	Estimate[a]	Std. error	Pr(>|t|)	R^2	AIC
EXP	0.12	0.01	0.00	0.97	−296.55
PL	0.72	0.03	0.00	0.97	−286.55
TPL	0.37	0.03	0.00	0.99	−355.29

[a] Estimated: estimated value to network; Std error: standard error; Pr(>|t|): model significance (values < 0.05 are considered significant); R^2: determination coefficient of the regression; AIC: Akaike information criterion.

in a total of 3850 syllables. Moreover, of all 170 analyzed individuals, 36 (21.5%) had multiple songs and, of those, 32 individuals had two songs. Four individuals had three songs. Song length (total number of syllables in a given song), on the other hand, varied from 7 to 48 ($\bar{x} = 18.5$, $SD = 5.9$), suggesting that some syllables were used considerably more often than others. When we analyzed *MQuant* for Zipf's law, we found that the frequency of syllable use distribution was best fit by a truncated power law, and only six syllables (34, 1, 10, 11, 20, and 7) account for nearly 50% of the recorded syllables (Table 1, Fig. 4). The L, D, and C of the syllable network were 1.695, 3, and 0.722, respectively, while L and C were significantly higher than those expected from random networks ($L_{null} = 1.633 \pm 0.001$, $D_{null} = 2.955 \pm 0.289$, $C_{null} = 0.602 \pm 0.001$, $P < 0.01$ in all comparisons). The index of small-word-ness S^{ws} had a value of 1.15, which indicates that the network can be regarded as small-word (Humphries and Gurney, 2008). Furthermore, the degree distribution of the *MSeq* also followed a distinct pattern than expected by the null model: 26 out of 30 points (86.6%) fell outside the null distribution generated by the random simulations. Both less connected and highly connected syllables had a higher frequency than expected by chance (Fig. 5).

The comparison of the level of fit of alternative models showed that the degree distribution of the *MSeq* was best fit by a truncated power law, both for the preceding syllable as subsequent syllable networks (Table 2, Fig. 6). There was also a significant association between the frequency of self- and out-degrees of each syllable ($r = 0.68$, $N = 42$, $P < 0.001$), indicating the more frequently associated with other syllables also tend to be more associated with
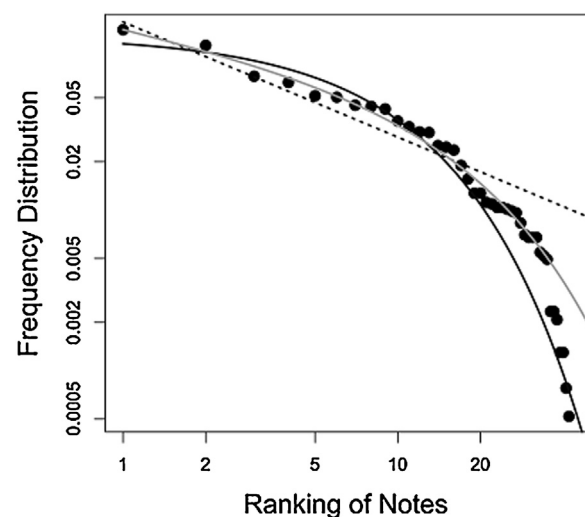

Fig. 4. Fit of the Zipf's Law in *MQuant* to models: exponential (black line), power law (dotted line) and truncated power law (gray line) in logarithmic scale. The best fit was obtained by the truncated power law model. It indicates that the frequency of syllable use initially decays as a power-law and exhibits a vertical truncation for the least frequent syllables. Also, the first six syllables (classified respectively as 34, 1, 10, 11, 20 and 7) correspond to 50% of the most frequently used syllables in all songs.
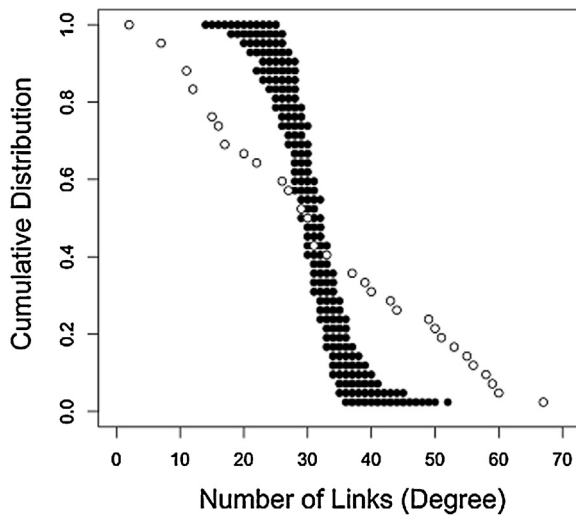
Fig. 5. Degree distribution of the matrix of notes association (*MSeq*). The white points correspond to degree distribution of MSeq and the black points represent the confidence interval of the null model. The curve of degree distribution fell out of the null space generated by the model. It points out for a higher frequency of syllables with low degree and a smaller frequency of highly connected syllables than expected by chance.
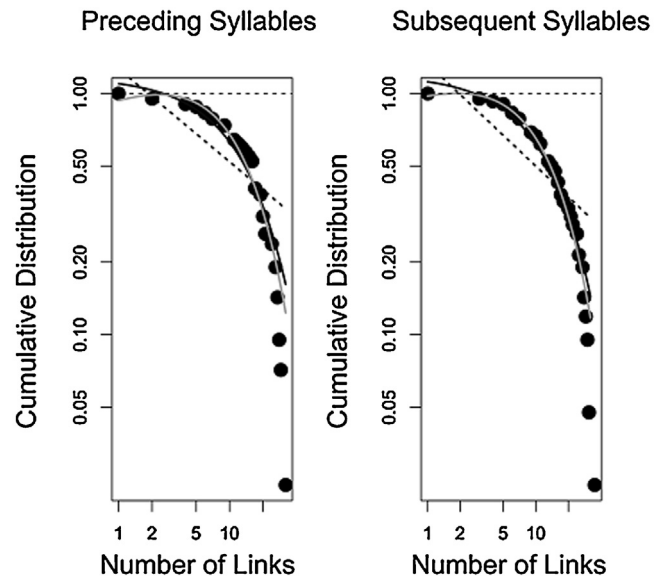

Fig. 6. Degree distributions for the syllable association network (*MSeq*) in logarithmic scale. The best fit was obtained by the truncated power law model for both preceding (in-degree) and subsequent syllables (out-degree). Exponential, power law and truncated power law fits are respectively represented by black, dotted and gray lines.

themselves. These results were further explored by comparing their frequencies with a line through the origin and with slope of one, which would indicate that interactions with themselves and interactions with other syllables were directly proportional. Although most syllables fall below this line, indicating preferential associations with other syllables, there were twelve types that show more self interactions than expected, corresponding to those that form phrases (Fig. 7).

Nestedness analyses based on the WINE method were consistent with the existence of substantial nestedness in both *MQuant* and *MSeq*, with estimated values being significantly higher than those obtained using the null models ($Z_{score} = 28.908$; $P < 0.01$ and $Z_{score} = 11.074$; $P < 0.01$), for *MQuant* and *MSeq* respectively. A nested pattern can be observed in the syllable association graphs, where the width of each vertex is drawn proportionately to the number of interactions between a given pair of syllables (Fig. 8). The most frequent syllable in the sample is the number 34, followed by 10, and 7.

4. Discussion

Our study involved a relatively broad sample of notes and, although we concede that it is far from a complete description of

Table 2
Relative fit of the *MSeq* degree distribution according to three models: Exp: Exponential; PL: Power Law and TPL: Truncated Power Law.

	Estimate[b]	Std. error	Pr(>|t|)	R^2	AIC
In-degree[a]					
E	0.06	0.00	0.00	0.97	−50.38
PL	0.38	0.06	0.00	0.82	−9.43
TPL	−0.20	0.05	0.00	0.97	−64.11
Out-degree[a]					
E	0.07	0.00	0.00	0.99	−79.02
PL	0.42	0.06	0.00	0.84	−14.11
TPL	−0.16	0.03	0.00	0.99	−103.17

[a] In-degree represents the fit of the subsequent syllables of the song; Out-degree, the fit of preceding notes.

[b] Estimate: estimated value to network; Std error: standard error; Pr(>|t|): model significance (values < 0.05 are considered significant); R^2: determination coefficient of the regression; AIC: Akaike information criterion.

the repertoire of the species, our coverage was sufficient to include a large sample of syllables and their associations. For instance, a study analyzing a comprehensive sample of the syllable repertoire of *Troglodytes aedon* (3574 songs of a single individual) showed that the repertoire in that species includes between 37 and 90 syllable types (Kroodsma, 1973). Tubaro (1990) characterized the repertoire of *T. aedon* in two populations in Argentina and also demonstrated similar values for repertoire size (38 different syllable types) and number of song types as those found in our study for *T. musculus.* Therefore, we believe that our sampling (from 170 individuals and 207 songs) is a representative estimate of the syllable repertoire of the species, at least at the geographical scale involved in our study.
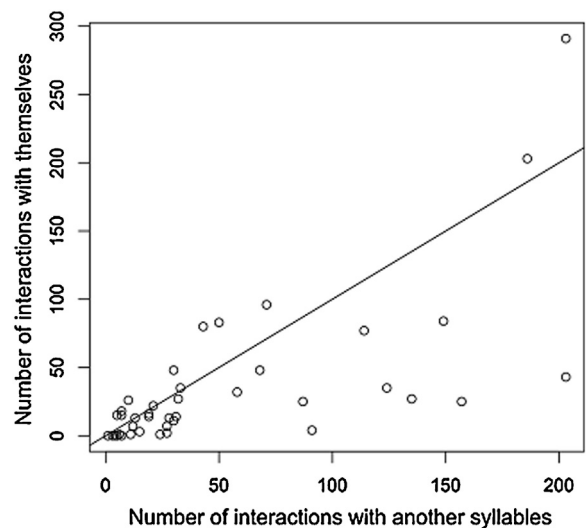

Fig. 7. Correlation between the number of interactions of a given syllable with itself and with other syllables ($r = 0.68$, $N = 42$, $P < 0.01$). The points above the line correspond to syllables that form the song phrases: 29, 22, 1, 6, 34, 5, 15, 2, 19, 24, 9, and 27. These syllables types tend to be repeated more frequently than they are followed by a different syllable.
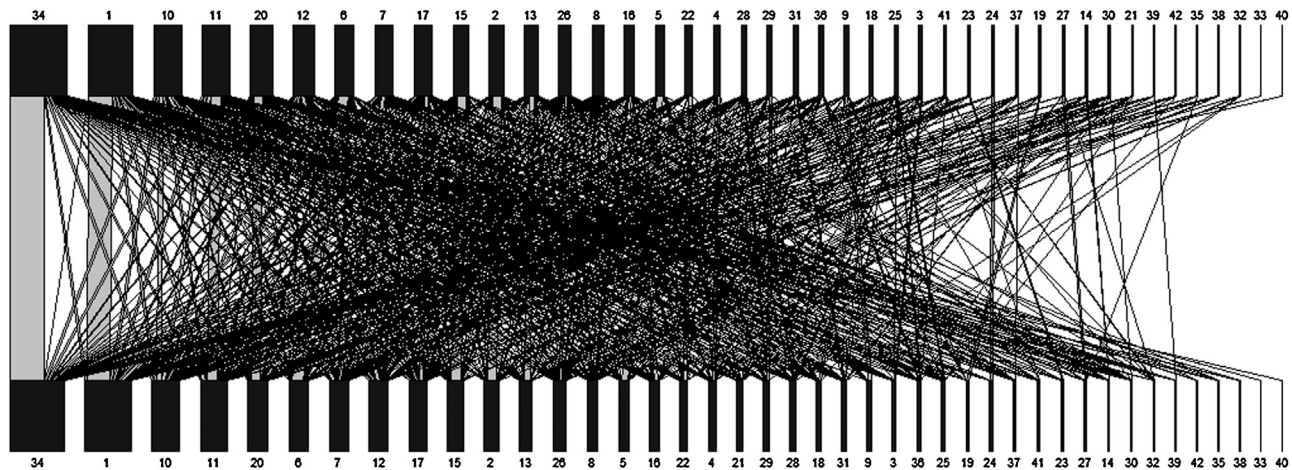
Fig. 8. Bipartite representation of *MSeq* network which depicts a nested pattern. Syllables with many connections interact with both highly and lowly connected syllables, as syllables with few connections tend to interact with the most connected ones. Upper and lower vertical lines represents preceding and subsequent syllables, respectively. They are ordered from the most connected syllables (left) to the less connected ones (right). The width of these lines is proportional to the number of interactions of each syllable. The longer vertical or diagonal lines in the middle represent the occurrence of an interaction between a pair of syllables.

The pattern of syllable associations' in *T. musculus* songs can be classified as a small-world network, given that the average distance was lower and the clustering coefficient was higher than expected by chance (Watts and Strogatz, 1998). Interestingly, similar patterns were observed in word association networks of several languages (Steyvers and Tenenbaum, 2005; Solé et al., 2010). Moreover, there are qualitative similarities in the degree distribution among syllables and words association networks, such as the large number of syllables with a low observed number of associations, whereas other syllables are used more than expected by chance (Fig. 5). On the other hand, the syllable degree distribution was best fit by a truncated power law, whereas word association networks are more commonly described by scale-free distributions. These networks also differ with respect to the slope of the degree distribution. Ferrer i Cancho et al. (2004) investigated several word association networks and found a slope of 2.2, whereas the slopes of the networks in the present study are considerably lower (Table 2: in-degree = −0.20 and out-degree = −0.16). It is possible that these discrepancies occur because syllable association networks are considerably smaller than word association networks, as well as the fact that avian songs must always maintain the goal of species recognition (see below). Nevertheless, the concordance in *C*, *L* and the general shape of the degree distributions might be interpreted as indicating common underlying mechanisms that could be universal for communication networks.

Briefer et al. (2010) showed that the frequency distribution of note use in the Eurasian Skylark *A. arvensis* deviates from a simple power law distribution expected from Zipf's law, following instead a curve reaching asymptotically a slope of zero. They interpreted this pattern as an indication that the syllable repertoire in this species can be a closed system due to the learning strategy based on imitation of young birds from their tutors, without the role of the improvisation. However, we interpret our results in the case of *T. musculus* in a slightly different manner. Rather, we believe that best fit of the distribution of degrees by a truncated power law function could have three implications: (i) a power law model could indicate that improvisation might be important in the song syntax, and that syllable composition in *T. musculus* is open to novel syllables (Catchpole, 1976); (ii) the final truncation suggests a possible constraint to the level of improvisation; and (iii) the introduction of new syllables might occur through errors during the process of copying the tutor by young males. A test of the latter hypothesis would require an experiment tutoring birds of different ages to

investigate if new notes (or rare notes) are product of improvisation or errors during the learning process.

We found substantial levels of nestedness in both *MQuant* and *MSeq*. Nestedness is a pervasive feature of several real-world networks (e.g. Bascompte et al., 2003; Guimarães et al., 2006; Ulrich et al., 2009). In a context of birdsong structure, a nested pattern implies that syllables that are highly connected tend to be associated with both highly-connected and less-connected syllables. On the other hand, syllables that connect with few others are most frequently associated with these highly connected ones. As a result, this network presents a topology in which low-connected syllables preferentially interact around a core of highly connected syllables. As for *MQuant*, the nested topology indicates that the most frequent syllables are associated with songs of both high and low number of syllable types, and songs with small number of syllable types tend to contain syllables from this core of most frequent ones. These patterns offer new insights in the study of bird song syntax because they allow for not only on the description of the frequency or the number of associations of each syllable, but also allow us to understand the characteristics of the nodes with which they are connected.

We believe that two mechanisms could be responsible for the nested organization of these networks. First, the simplest explanation is that a nested topology could be an epiphenomenon of the fact that syllables differ considerably in their frequency of use. As a consequence, rare syllables would simply tend to be associated by chance alone. Alternatively, there could be a syntactic value in a nested topology if more frequently used syllables are important for species recognition, whereas rare ones would be more important for female preference. If that is the case, there should be a trade-off between recognition and preference, with common syllables forming association cores and being interspersed with improvisation using rare syllables. This is also supported by the observation that shorter songs always maintain the most commonly used syllables. Although these are not mutually exclusive mechanisms, specific studies focusing on the pattern of syllables succession within single males could be particularly useful to assess their relative contribution.

The similarities between the topological properties of the syllable association networks recorded in our study for the house wren, as well as the results of other studies on avian communication, suggests the possibility of general principles underlying avian syllable association networks. For instance, the song network in the California Trasher (*Toxostoma redivivum*) also followed

a small-world pattern and a non bell-shaped degree distribution $P(k)$, despite large individual differences in the sizes of their phrase repertoires and their specific phrase types (Sasahara et al., 2012). Despite using simpler methods that do not involve explicitly reconstructing the topology of syllable association networks, previous studies investigating syntax in oscine birds had already found evidence for predictability in song organization. For instance, both in sedge warbler *Acrocephalus schoenobaenus* (Catchpole, 1976) and in the willow warbler *Phylloscopus trochilus* (Gill and Slater, 2000), syllable organization is clearly nonrandom, such that their probabilities of occurrence depends at least on the immediately preceding syllable. Moreover, most willow warbler songs start with the same element and there after elements are produced in a highly predictable way.

The network analysis of syllables associations is still a poorly explored view into the study of song evolution. In the present study we demonstrate that this approach can provide several interesting insights in relation to traditional methods, allowing for a more precise description of songs with complex syntax, especially in the case of oscine birds. Moreover, it would be of particular interest to establish the extent of the parallels between syllables and word association networks, such as the relationship between syllable use and their rate of evolutionary change (Pagel et al., 2007).

Acknowledgments

References

Abe, K., Watanabe, D., 2011. Songbirds posses the spontaneous ability to discriminate syntactic rules. Nature Neuroscience 17, 1067–1074.

Albert, R., Barabási, A.L., 2002. Statistical mechanics of complex networks. Reviews of Modern Physics 74, 47–97.

Amaral, L.A.N., Scala, A., Barthelemy, M., Stanley, H.E., 2000. Classes of small-world networks. Proceedings of the National Academy of Sciences of the United States of America 97, 11149–11152.

Balaban, E., 1988. Bird song syntax: learned intraspecific variation is meaningful. Proceedings of the National Academy of Sciences of the United States of America 85, 3657–3660.

Bascompte, J., Jordano, P., Melián, C.J., Olesen, J.M., 2003. The nested assembly of plant-animal mutualistic networks. Proceedings of the National Academy of Sciences of the United States of America 100, 9383–9387.

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

Botero, C.A., Mudge, A.E., Koltz, A.M., Hochachka, W.M., Veherencamp, S.L., 2008. How reliable are the methods for estimating repertoire size? Ethology 114, 1227–1238.

Bouttier, J., Di Francesco, P., Guitter, E., 2003. Geodesic distance in planar graphs. Nuclear Physics B 663, 535–567.

Brewer, D., 2001. Wrens, Dippers and Thrashers. Yale University Press, New Haven/London.

Briefer, E., Aubin, T., Lehongre, K., Rybak, F., 2008. How to identify dear-enemies: the group signature in the complex song of the skylark *Alauda arvensis*. Journal of Experimental Biology 211, 317–326.

Briefer, E., Osiejuk, T.S., Rybaka, F., Aubin, T., 2010. Are bird song complexity and song sharing shaped by habitat structure? An information theory and statistical approach. Journal of Theoretical Biology 262, 151–164.

Butts, C.T., 2010. SNA: Tools for Social Network Analysis R Package Version 2.2-0. http://cran.r-project.org/package=sna

Catchpole, C.K., 1976. Temporal and sequential organization of song in the sedge warbler (*Acrocephalus schoenobaenus*). Behaviour 59, 226–246.

Catchpole, C.K., Slater, P.J.B., 2008. Bird Song: Biological Themes and Variation. Cambridge University Press, Cambridge.

Cavalli-Sforza, L.L., Piazza, A., Menozzi, P., Mountain, J., 1988. Reconstruction of human evolution: bringing together genetic, archaeological, and linguistic data. Proceedings of the National Academy of Sciences of the United States of America 85, 6002–6006.

Corominas-Murtra, B., Valverde, S., Solé, R., 2009. The ontogeny of scale-free syntax networks: phase transitions in early language acquisition. Advances in Complex Systems 12, 371–392.

Csardi, G., Nepusz, T., 2006.The igraph software package for complex network research. InterJournal: Complex Systems, 1695. http://igraph.sf.net/

Dormann, C.F., Gruber, B., Fründ, J., 2008. Introducing the bipartite Package: analyzing ecological networks. R News 8, 8–11.

Dormann, C.F., Fründ, J., Blüthgen, N., Gruber, B., 2009. Indices, graphs and null models: analyzing bipartite ecological networks. The Open Ecology Journal 2, 7–24.

Doupe, A.J., Kuhl, P.K., 1999. Birdsong and human speech: common themes and mechanisms. Annual Review of Neuroscience 22, 567–631.

Erdös, P., Rényi, A., 1959. On random graphs. Publicationes Mathematicae 6, 290–297.

Ferrer i Cancho, R., Solé, R.V., 2001. The small world of human language. Proceedings of the Royal Society B: Biological Sciences 268, 2261–2265.

Ferrer i Cancho, R., Solé, R., Köhler, R., 2004. Patterns in syntactic dependency networks. Physical Review E 69, 1–8.

Galeano, J., Pastor, J.M., Iriondo, J.M., 2009. Weighted-Interaction Nestedness Estimator (WINE): a new estimator to calculate over frequency matrices. Environmental Modeling Software 24, 1342–1346.

Gill, D., Slater, P.J.B., 2000. Song organization and singing patterns of the willow warbler, *Phylloscopus trochilus*. Behaviour 137, 759–782.

Guimarães, P.R., Rico-Gray, V., Reis, S.F., Thompson, J.N., 2006. Asymmetries in specialization in ant-plant mutualistic networks. Proceedings of the Royal Society B: Biological Sciences 273, 2041–2047.

Holland, J., Dabelsteen, T., Paris, A., 2000. Coding in the song of the wren: importance of rhythmicity, syntax and element structure. Animal Behaviour 60, 463–470.

Humphries, M.D., Gurney, K., 2008. Network 'Small-World-Ness': a quantitative method for determining canonical network equivalence. PLoS ONE 3, 2051.

Kendeigh, S.C.,1941. Territorial and mating behavior of the house wren. In: Illinois Biological Monographs XVIII. The University of Illinois, Urbana.

Kreutzer, M., Brémond, J.C., 1986. Les effects additifs de la syntax et de la forme des syllabes lors de la reconnaissance spécifique chez le troglodyte (*Troglodytes troglodytes*). Canadian Journal of Zoology 64, 1241–1244.

Kroodsma, D.E., 1973. Coexistence of bewick's wrens and house wrens in Oregon. The Auk 90, 341–352.

Kroodsma, D., 2005. The Singing Life of Birds: The Art and Science of Listening to Birdsong. Houghton Mifflin Harcourt, New York.

Lipkind, D., Marcus, G.F., Bemis, D.K., Sasahara, K., Jacoby, N., Takahasi, M., Suzuki, K., Feher, O., Ravbar, P., Okanoya, K., Tchernichovski, O., 2013. Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. Nature 498, 104–108.

Nowicki, S., Nelson, D.A., 1990. Defining natural categories in acoustic signals: comparison of three methods applied to 'Chick-a-dee' call notes. Ethology 86, 89–101.

Pagel, M., Atkinson, Q., Meade, A., 2007. Frequency of word-use predicts rates of lexical evolution throughout Indo-european history. Nature 449, 717–720.

Peters, S.S., Searcy, W.A., Marler, P., 1980. Species song discrimination in choice experiments with territorial male swamp and song sparrows. Animal Behaviour 28, 393–404.

Podos, J., Peters, S., Rudnicky, T., Marler, P., Nowicki, S., 1992. The organization of song repertoires in song sparrows: themes and variations. Ethology 90, 89–106.

Podos, J., Warren, P.S., 2007. The evolution of geographic variation in birdsong. Advances in the Study of Behavior 37, 403–458.

Proulx, S.R., Promislow, D.E.L., Phillips, P.C., 2005. Network thinking in ecology and evolution. Trends in Ecology and Evolution 20, 345–353.

Sasahara, K., Cody, M.L., Cohen, D., Taylor, C.E., 2012. Structural design principles of complex bird songs: a network-based approach. PLoS ONE 7, 1–9.

Searcy, W., McArthur, P., Peters, S., Marler, P., 1981. Responses of male song and swamp sparrow to neighbor, stranger and self song. Behaviour 77, 152–163.

Sick, H., 1997. Ornitologia Brasileira. Nova Fronteira, São Paulo.

Solé, R., 2005. Syntax for free? Nature 434, 289.

Solé, R., Corominas-Murtra, B., Valverde, S., Steels, L., 2010. Language networks: their structure, function, and evolution. Complexity 15, 20–26.

Steyvers, M., Tenenbaum, J.B., 2005. The large-scale structure of semantic networks: statistical analyses and a model of semantic growth. Cognitive Science 29, 41–78.

Tubaro, P.L., 1990. Song description of the House Wren (*Troglodytes aedon*) in two populations of eastern Argentina, and some indirect evidences of imitative vocal learning. El Hornero 13, 111–116.

Ulrich, W., Almeida-Neto, M., Gotelli, N.J., 2009. A consumer's guide to nestedness analysis. Oikos 118, 3–17.

Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of small-world networks. Nature 393, 440–442.