Do penguins’ vocal sequences conform to linguistic laws?

Published version:

DOI:10.1098/rsbl.2019.0589

BIOLOGY
LETTERS

Do penguins' vocal sequences conform to linguistic laws?

Journal:	*Biology Letters*
Manuscript ID	RSBL-2019-0589.R2
Article Type:	Research
Date Submitted by the Author:	n/a
Complete List of Authors:	Favaro, Livio; University of Torino, Life Sciences and Systems Biology; University of Lyon/Saint-Etienne, Equipe de Neuro-Ethologie Sensorielle ENES/CRNL, CNRS UMR5292, INSERM UMR_S 1028 Gamba, Marco; University of Torino, Life Sciences and Systems Biology Cresta, Eleonora; University of Rome La Sapienza, Department of Environmental Biology Fumagalli, Elena; University of Milan, Department of Environmental Science and Policy Bandoli, Francesca; Giardino Zoologico di Pistoia Pilenga, Cristina; Zoomarine Italia Isaja, Valentina; Zoom Torino Mathevon, Nicolas; University of Lyon/Saint-Etienne, Equipe de Neuro-Ethologie Sensorielle ENES/CRNL, CNRS UMR5292, INSERM UMR_S 1028 Reby, David; University of Lyon/Saint-Etienne, Equipe de Neuro-Ethologie Sensorielle ENES/CRNL, CNRS UMR5292, INSERM UMR_S 1028
Subject:	Behaviour < BIOLOGY
Categories:	Animal Behaviour
Keywords:	Acoustic sequences, Bioacoustics, Compression, Information theory, Seabirds

Author-supplied statements

Relevant information will appear here if provided.

Ethics

Does your article include research that required ethical approval or permits?:
This article does not present research with ethical considerations

Statement (if applicable):
Our study complies with all regulations for animal care in Italy. According to Italian laws, no specific permissions were required for collecting non-invasive passive acoustic recordings of penguins. All experimental procedures are also in accordance with French national guidelines, permits and regulations regarding animal care and experimental use (Approval nÂ°D42-218-0901, ENES lab agreement, Direction DÃ©partementale de la Protection des Populations, PrÃ©fecture du RhÃ´ne).

Data

It is a condition of publication that data, code and materials supporting your paper are made publicly available. Does your paper present new data?:
Yes

Statement (if applicable):
Additional information is available in the Electronic Supplementary Material.

Conflict of interest

I/We declare we have no competing interests

Statement (if applicable):
CUST_STATE_CONFLICT :No data available.

Authors' contributions

This paper has multiple authors and our individual contributions were as below

Statement (if applicable):
LF, MG, and DR conceived/designed the study. CP, FB, VI, and LF coordinated the fieldwork. EC and EF collected the acoustic recordings, visually inspected the spectrograms, and labelled the vocal sequences. MG wrote the Praat scripts for the acoustic analysis. LF and MG carried out the statistical analyses. LF, DR, and NM made a substantial contribution to the interpretation of the data. LF took the lead in writing the manuscript. DR, MG, and NM helped draft the manuscript. EC, CP, EF, FB, and VI revised the manuscript critically for important intellectual content. All authors gave final approval for publication and agree to be held accountable for the work performed therein.

Do penguins' vocal sequences conform to linguistic laws?

Livio Favaro[1,2], Marco Gamba[2], Eleonora Cresta[3], Elena Fumagalli[4], Francesca Bandoli[5], Cristina Pilenga[6], Valentina Isaja[7], Nicolas Mathevon[1], David Reby[1]

[1] Equipe de Neuro-Ethologie Sensorielle ENES/CRNL, University of Lyon/Saint-Etienne, CNRS UMR5292, INSERM UMR_S 1028, Saint-Etienne, France
[2] Department of Life Sciences and Systems Biology, University of Turin, Turin, Italy
[3] Department of Environmental Biology, Sapienza University of Rome, Rome, Italy
[4] Department of Environmental Science and Policy, University of Milan, Milan, Italy
[5] Giardino Zoologico di Pistoia, Pistoia, Italy
[6] Zoomarine Italia, Torvaianica-Pomezia, Rome, Italy
[7] Zoom Torino, Piscina, Turin, Italy

Author for correspondence:
Livio Favaro
email: livio.favaro@unito.it

Abstract

Information compression is a general principle of human language: the most frequent words are shorter in length (Zipf's law of brevity) and the duration of constituents decreases as the size of the linguistic construct increases (Menzerath-Altmann law). Vocal sequences of non-human primates have been shown to conform to both these laws, suggesting information compression might be a more general principle. Here we investigated whether display songs of the African penguin, which mediate recognition, intersexual mate choice, and territorial defence, conform with these laws. Display songs are long, loud sequences combining three types of syllables. We found that the shortest type of syllable was the most frequent (with the shortest syllable being repeated stereotypically, potentially favouring signal redundancy in crowded environment). We also found that the average duration of the song's constituents was negatively correlated with the size of the song (a consequence of increasing the relative number of the shortest syllable type, rather than reducing the duration across all syllable types, thus preserving the communication of size related information in the duration of the longest syllable type). Our results provide the first evidence for conformity to Zipf's and Menzerath-Altmann laws in the vocal sequences of a non-primate species, indicating that these laws can co-exist with selection pressures specific to the species' ecology.

Keywords

Acoustic sequences, Bioacoustics, Compression, Information theory, Seabirds.

1. Introduction

Quantitative linguistics is directing increasing research effort toward the identification and study of universal statistical patterns, known as "linguistic laws", occurring in human, animal, and artificial communication systems [1, 2, 3]. Among these, of particular importance are Zipf's Law of Brevity and the Menzerath-Altmann Law. Zipf's law postulates that more frequent elements tend to be compressed [4] and has been found to apply in all tested human languages, where the words used most often are shorter [5, 6]. Similarly, the Menzerath-Altmann law postulates that, in human language, the sizes of the constituents (e.g. phonemes) of a construction (e.g. morpheme) decrease with increasing size of the construction [7]. Overall, these two principles suggest that, in human vocal communication, the maximisation of coding efficiency and minimization of code length act as selective pressures to compress the elements supporting information.

Comparative studies have found evidence for the occurrence of the Zipf's Law of Brevity in vocal sequences of non-human primates [8,9], and some authors suggested that, in vocal sequences, vocal compression might be a general principle beyond the human language [10]. However, a lack of conformity to the law has been found in the vocal repertoire of golden-backed uakaris (*Cacajao melanocephalus*) [9]. More recently, the size of vocal sequences uttered by wild gelada males (*Theropithecus gelada*) was found to be negatively correlated with the duration of the calls constituting the sequence, providing the first indication of Menzerath-Altmann Law conformity in a non-human primate [2]. Evidence of conformity to Menzerath-Altmann Law comes also from Chimpanzee pant hoots [11]. However, more *taxa* remain to be investigated, especially beyond primates, to develop a broader understanding of the occurrence and significance of such statistical patterns in animal vocalisations.

Sequences of acoustic elements are ubiquitous in animal vocal communication systems [12]. Several mammals, such as bats, carnivores, cetaceans, primates, and rodents produce sequences of acoustic units to convey information on species identity, predators, and resources [13-15]. Nevertheless, most research in this area has focused on learned bird songs. Bird songs are series of discrete acoustic elements, usually referred as "syllables", separated by silent intervals [16], that mediate mate choice, social recognition and territorial defence. Accordingly, the sequence can encode a variety of information ranging from the individual identity of the emitter to its physical quality and social rank [17].

The use of vocal sequences is also common in seabirds [18], despite the absence of evidence for vocal production learning. It is particularly prevalent in penguins [19], a family of non-flying seabirds that diverged from the main bird lineage approximately 71 million years ago [20] and have evolved a complex vocal repertoire comprising both calls and display songs [19, 21-23]. Display songs are made of sequences of identical and redundant acoustic units, whose spectral envelope encode cues to the individual identity of the emitter [24]. However, as an exception to this general rule, ecstatic display songs (hereafter EDS) of the African penguin (*Spheniscus demersus*) are made of sequences of three distinct and discrete acoustic elements (hereafter syllables). EDS are typically composed of an initial series of short (0.18 ± 0.05 s) “type-A” syllables followed with by a longer (1.14 ± 0.33) type-B syllable, both produced on exhalation [21]. Following the emission of each type-A or type-B, penguins can produce a further type of syllable of intermediate duration (0.38 ± 12) while inhaling air (type-C) [21]. This resulting unusual pattern of inhalation/exhalation calls resembles a donkey bray hence this species' vernacular name: the “jackass penguin” [19]. Moreover, in *Spheniscus* penguins, the duration and fundamental frequency of the type-B syllables correlates with the body size of the emitter, suggesting that the spectrotemporal features of this syllable type may function to advertise respiratory capacity and thus quality in intersexual mate choice or intrasexual competition contexts [25].

Our study investigated whether EDS of the African penguin conform to Zipf's Law of Brevity and the Menzerath-Altmann Law. In particular, we tested whether (1) the length of the syllables decreases with their frequency of use, and (2) the average duration of the constituting syllables decreases as the size of the vocal sequence increases.

2. Materials and Methods

2.1 Acoustic recordings

We recorded 590 ecstatic display songs (Figure 1) from 28 adult African penguins (mean per bird = 21.07, Standard Error = 3.05) belonging to three different *ex situ* colonies in Italian zoos. Detailed information for each colony as well as the contribution of each bird recorded in this study is provided in the Electronic Supplementary Material. We collected vocalisations during the breeding periods in 2016 and 2017, using a focal animal sampling method [26]. We used a RØDE NTG-2 Super-Cardioid microphone (frequency response 20 Hz to 20 kHz, sensitivity -36dB +/- 2 dB re 1 V/Pa at 1 kHz, max SPL 131dB) handled using on a RØDE PG2 Pistol Grip and placed 5 to 10 meters from the vocalising individuals. The microphone output signal was digitalised using a TASCAM DR-680 or TASCAM DR-40 professional recorders at 44.1 kHz sampling rate and saved into an internal SD memory card in WAV format (16-bit amplitude resolution).

2.2 Analysis of acoustic sequences

Spectrograms of EDS were visually inspected in Praat v. 6.0.43 [27] by EF and EC. Using a one-layer TextGrid object, we assigned each vocal element to a syllable type based on its spectrotemporal features, as previously described in the literature [19,21]. Overall, we identified 7,495 syllables (type-A = 4,440; type-B = 1,298; type-C = 1,757). Subsequently, 5% of the syllables were randomly chosen to be inspected by a third observer (LF). The inter-observer reliability was extremely high (Cohen's Kappa coefficient = 0.99).

Using a custom-built script in Praat, we calculated the number of syllables constituting each EDS (sequence size), the duration of the vocal sequence, the duration of each single syllable, and the proportion of each syllable type (number of syllables of a given type divided by the total number of syllables).

2.3 Statistical analysis

To investigate whether the type of syllable has an effect on the frequency of use, we built a Linear Mixed effect Models (LMM) using the *lmer* package [28] in R v. 3.5.2 [29]. We included the syllable proportion in the EDS as the response variable and the type of syllable as fixed factor. The individual, colony and size of the vocal sequence were included as random effects to account for repeated measurements. We looked at a qqplot and the distribution of the residuals plotted against the fitted values to check for the assumptions that the model residuals were homogeneous and normally distributed. Using a likelihood ratio test [30], we then tested the significance of a "full model" comprising both the fixed and random factors against a "null model" comprising the random factor exclusively [31]. The alpha value for the predictor was calculated using the R-function "drop1" [32]. Moreover, using the "multcomp" package, we performed a Tukey post hoc test for which we reported estimate, z- and *p-values*. Finally, to confirm that the resulting most frequently used syllable type corresponded to the shortest in duration, we built an additional LMM model using the duration of the syllable as the response variable, the type of syllable as fixed factor, and the individual, colony, and EDS as random effects.

Furthermore, we explored the relationship between the number of the syllables constituting the EDS (*i.e.* sequence size) and their duration with a LMM where the duration of the syllables was used as the response variable and the total number of the syllables as fixed factor. The individual, and colony were included as random effects. Since an effect of the predictor was detected, we further investigated whether this was due to a change in the type of syllables constituting the sequences or variation in the duration of specific syllable types. Accordingly, we ran six LMMs (two for each syllable type), using the total number of the syllables in the EDS as fixed factor and, as the response variable, the number of syllables of that given type and their duration, respectively. For all these models, the individual and colony were included as random effects to account for repeated measurements.

3. Results

We found that in display songs of the African penguin, the proportion of type-A syllables (mean ± SD = 0.58 ± 0.16) is far greater than those of the type-C (0.24 ± 0.09) and the long type-B (0.18 ± 0.08). Accordingly, the LMM models detected a strong effect of the type of syllable on its frequency of use (LMM full vs. null: $\chi 2$ = 2008.772, d*f* = 1, *p* < 0.001; Tukey post hoc: type-A *vs* type-B estimate = -0.400 ± 0.007, *z* = -56.910*, p* = < 0.001; type-A *vs* type-C estimate = -0.342 ± 0.008, *z* = -48.348*, p* = < 0.001; type-B *vs* type-C estimate = 0.058 ± 0.007, *z* = 8.241*, p* = < 0.001). Importantly, our results also confirmed the presence of significant differences in the duration of the three syllables (LMM full vs. null: $\chi 2$ = 14683.85, *df* = 2, *p* < 0.001), with the shortest elements being the type-A (mean ± SD = 0.210 ± 0.053 s), the longest the type-B (1.163 ± 0.304 s), and the type-C of intermediate duration (0.304 ± 0.125 s).

We found that the number of syllables of the EDS (sequence size) has an effect on the duration of the syllables (LMM full vs. null: $\chi 2$ = 5.843, *df* = 1, *p* < 0.001). However, we found that such effect is due to a variation in the type of syllables constituting the song, with sequences made of more elements having a far greater number of the short type-A (LMM full vs. null: $\chi 2$ = 705.48, *df* = 1, *p* < 0.001; estimate = 0.794 ± 0.012; Figure 2a), increased number of inhaling type-C (LMM full vs. null: $\chi 2$ = 133.79, *df* = 1, *p* < 0.001, estimate = 0.178 ± 0.015; Figure 2a), and only slightly more long type-B units (LMM full vs. null: $\chi 2$ = 10.27, *df* = 1, *p* < 0.001, estimate = 0.032 ± 0.010; Figure 2a). Moreover, we found that neither type-A (LMM full vs. null: $\chi 2$ = 0.073, *df* = 1, *p* = 0.786; Figure 2b) nor type-B (LMM full vs. null: $\chi 2$ = 0.511, *df* = 1, *p* = 0.474; Figure 2b) syllables shorten in duration with respect to the number of elements constituting the sequence, although an effect was detected for the inhaling type-C (LMM full vs. null: $\chi 2$ = 32.498, *df* = 1, *p* < 0.001, estimate = -0.005 ± 0.001; Figure 2b).

4. Discussion

Our results demonstrate that ecstatic display songs of the African penguin follow Zipf's Law of Brevity and the Menzerath-Altmann Law. This is the first compelling evidence for conformity to linguistic laws in vocal sequences of a non-primate species. As predicted, we found that the duration of the syllables was inversely correlated with the frequency of occurrence. In particular, we found the short type-A the most frequent and the long type-B the least frequent in the ecstatic display songs. This conforms to the universal principle of compression, which states that the most frequently used elements in communication systems tend to be the shortest [33]. However, it is important to note that, in this case, the conformity has only been tested in a subset of the vocal repertoire (i.e. EDS), where the elements are repeated stereotypically and consecutively.

We also found that the sequences appear to reflect the Menzerath-Altmann principle, whereby larger constructs (vocal sequences) are made of smaller-size constituents (syllables) [34]. Indeed, the number of syllables constituting the EDS (i.e. sequence size) was negatively correlated with the average duration of their constituting syllables. However, we found that larger sequences (with more acoustic elements) were not achieved by a global reduction of the duration of all syllable types, but rather by a change in their composition: larger sequences were achieved by including a larger number of the short type-A syllables (of normal duration), rather than by shortening the duration of syllables across all syllable types. Similar conformity has been observed in vocal sequences of male geladas (*Theropithecus gelada*), where the emission of sequences with a greater number of calls is also achieved by increasing the number of shorter call types, probably reflecting energetic/breathing constraints on vocal production [2].

Analogies between the frequency of use and duration of acoustic elements in human language and animal vocalisations are highly debated [8-9, 35]. In the absence of lexical syntax [36], the duration of non-human animal vocalisations is likely to mainly result from selective pressures favouring the encoding and transmission of indexical information according to species-specific ecological needs. For example, long and high-amplitude vocalisations are efficient for long-distance communication [37], while short and low-amplitude calls are likely to occur in behavioural contexts where individuals are close to each other since they attenuate shortly in the environment [38]. In this respect penguins provide an interesting model, as they breed in large and noisy colonies, where recognition occur at a close distance and is mediated by display songs [24]. Here, the presence of redundant acoustic elements (syllables) has been attributed to the Shannon information theory [39], where multiple (identical) units improve the probability of receiving a message in a noisy environment [40]. In the non-nesting king penguin (*Aptenodytes patagonicus*), which breeds in sub-Antarctic islands, the playback of one-third of a single syllable is sufficient to allow individual recognition [24], which indeed shows that individual identity information is redundant both between and within syllables. However, African penguins have completely different habits and breed in temperate areas with a far less constraining environment, where wind blows do not dramatically affect the signal-to-noise ratio and where visual cues provided by the nest assist and support the vocal recognition in dense colonies [41]. In such contrasting conditions, removing the redundant information encoded within syllables is likely to be favoured by natural selection, as minimising the length of the elements can be assumed to reduce the energetic cost of signal production. However, repeating the same syllable multiple times guarantee the information redundancy within the sequence and minimises the chance of sending the information overlapped with other individuals vocalising at the same time.

Banded penguins are nesting species with an unusual colonial lifestyle, in that they are highly territorial and build nests in underground burrows which they excavate themselves under rocks, bushes or in guano [41]. Consistent with the breeding ecology of these species, in addition to individual recognition, the EDS also play a key role in mediating territorial defence and mate choice [19, 41]. Favaro et al. [25] showed that the duration of the long type-B syllables is an "honest signal" that correlates with body dimension of the emitter in the closely related Humboldt (*Spheniscus humboldti*) and Magellanic (*S. magellanicus*) penguins, potentially reflecting lungs and aerial sacs capacity. In the African penguin, the duration of the type-B syllable might therefore be subject to similar sexual selection pressures for the communication of respiratory capacity, preventing the shortening of this acoustic unit. However, we found that type-C syllables tend to vary in duration with respect to the number of acoustic elements of the song. We suggest that the duration of this type of syllable reflect the breathing constraints following the emission of the preceding units. Indeed, preliminary analyses indicate that type-C syllables following type-A syllables are shorter (mean ± SD = 0.19 ± 0.06 s) than those following the type-B (0.37 ± 0.11 s).

As such an increase in the number of type-A syllables would result in an overall shortening of the duration of type-C syllables.

To conclude, our findings provide the first evidence for conformity to Zipf's and Menzerath-Altmann linguistic laws in vocal sequences of the African penguin. We suggest that relationships between element duration, frequency of use, and song size are mainly a consequence of vocal production constraints interacting with selective pressures for intersexual mate choice and territorial defence in dense colonies. Importantly, our results suggest for the first time that information compression can coexist with other sources of selection in a non-primate species with a small and relatively fixed vocal system.

References

[1] Torre IG. Luque B, Lacasa L, Miramontes O, Hernandez-Fernandez A. 2017 Emergence of linguistic laws in human voice. *Sci. Rep.* **7**, 43862. (doi: 10.1038/srep43862)

[2] Gustison ML, Semple S, Ferrer-i-Cancho R, Bergman TJ. 2016 Gelada vocal sequences follow Menzerath's linguistic law. *Proc. Natl Acad. Sci. USA* **20**, 2750-2758 (doi: 10.1073/pnas.1522072113)

[3] Heesen R, Hobaiter C, Ferrer-i-Cancho R, Semple S. 2019 Linguistic laws in chimpanzee gestural communication. *Proc. R. Soc. B* **286**, 20182900. (http://dx.doi.org/10.1098/rspb.2018.2900)

[4] Zipf GK. 1949 *Human behaviour and the principle of least effort*. Cambridge, MA: Addison-Wesley.

[5] Piantadosi ST, Tily H, Gibson E. 2011. Word lengths are optimized for efficient communication. *Proc. Natl. Acad. Sci. USA* **108**, 3526–3529. (doi: 10.1073/pnas.1012551108)

[6] Bentz C, Ferrer-i-Cancho R. 2015 Zipf's law of abbreviation as a language universal. Capturing Phylogenetic Algorithms for Linguistics. Lorentz Center Workshop, Leiden, October 2015.

[7] Altmann G. 1980 Prolegomena to Menzerath's law. *Glottometrika* **2**, 1-10.

[8] Semple S, Hsu MJ, Agoramoorthy G. 2010 Efficiency of coding in macaque vocal communication. *Biol. Lett.* **6,** 469-471. (doi: 10.1098/rsbl.2009.1062)

[9] Bezerra BM, Souto AS, Radford AN, Jones G. 2010 Brevity is not always a virtue in primate communication. *Biol. Lett.* **7**, 23–25. (doi: 10.1098/rsbl.2010.0455)

[10] Ferrer-i-Cancho R, Hernández-Fernández A, Lusseau D, Agoramoorthy G, Hsu MJ, Semple S. 2013 Compression as a universal principle of animal behavior. *Cogn. Sci*. **37**, 1565–1578. (doi: 10.1111/cogs.12061)

[11] Fedurek P, Zuberbühler K, Semple S. 2017 Trade-offs in the production of animal vocal sequences: insights from the structure of wild chimpanzee pant hoots. *Front. Zool.* **14**, 50. (doi:10.1186/s12983-017-0235-8)

[12] Kershenbaum A, Blumstein DT, Roch MA, Akçay C, Backus G, Bee MA, Bohn K, Cao Y, Carter G, Cäsar C, et al. 2016 Acoustic sequences in non-human animals: a tutorial review and prospectus. *Biol. Rev.* **91**, 13-52. (doi: 10.1111/brv.12160)

[13] Gamba M, Favaro L,Torti V, Sorrentino V, Giacoma C. 2011 Vocal tract flexibility and variation in the vocal output in wild indris. *Bioacoustics* **20**, 251–266. (doi: 10.1080/09524622.2011.9753649)

[14] Garland EC, Goldizen AW, Rekdahl ML, Constantine R, Garrigue C, Hauser ND, Poole MM, Robbins J, Noad MJ. 2011 Dynamic horizontal cultural transmission of Humpback whale song at the ocean basin scale. *Curr. Biol.* **21,** 687-691. (doi: 10.1016/j.cub.2011.03.019)

[15] Hechavarría JC, Beetz MJ, Macias S, Kössl M. 2016 Distress vocalization sequences broadcasted by bats carry redundant information. *J. Comp. Physiol. A* **202**, 503–515. (doi: 10.1007/s00359-016-1099-7)

[16] Catchpole C, Slater P. 2008 *Bird song biological themes and variations*. Cambridge, UK: Cambridge University Press.

[17] Holveck M, de Castro ACV, Lachlan RF, ten Cate C, Riebel K. 2008 Accuracy of song syntax learning and singing consistency signal early condition in zebra finches. *Behav. Ecol.* **19**, 1267–1281. (doi: 10.1093/beheco/arn078)

[18] Hardouin LA, Thompson R, Stenning M, Reby D. 2014 Anatomical bases of sex and size-related acoustic variation in herring gull alarm calls. *J. Avian Biol.* **45**, 157-166. (doi: 10.1111/j.1600-048X.2013.00144.x)

[19] Jouventin P. 1982 Visual and vocal signals in penguins, their evolution and adaptive characters. *Adv. Ethol.* **58**, 3-148. (doi: 10.1002/iroh.3510680523)

[20] Baker AJ, Pereira SL, Haddrath OP, Edge KA. 2006 Multiple gene evidence for expansion of extant penguins out of Antarctica due to global cooling. *Proc. R. Soc. B*. **273**, 11-17. (doi: 10.1098/rspb.2005.3260)

[21] Favaro L, Ozella L, Pessani D. 2014 The vocal repertoire of the African penguin (*Spheniscus demersus*): Structure and function of calls. *PLoS ONE* **9(7)**, e103460. (doi: 10.1371/journal.pone.0103460)

[22] Favaro L, Gamba M, Alfieri C, Pessani D, McElligott AG. 2015 Vocal individuality cues in the African penguin (*Spheniscus demersus*): a source-filter theory approach. *Sci. Rep.* **5**, 17255. (doi: 10.1038/srep17255)

[23] Favaro L, Gili C, Da Rugna C, Gnone G, Fissore C, Sanchez D, Mcelligott AG, Gamba M, Pessani D. 2016 Vocal individuality and species divergence in the contact calls of banded penguins. *Behav. Process.* **128**, 83-88. (doi: 10.1016/j.beproc.2016.04.010)

[24] Aubin T, Jouventin P. 2002 How to vocally identify kin in a crowd: The penguin model. *Adv. Study Beh.* **31**, 243-277. (doi: 10.1016/S0065-3454(02)80010-9)

[25] Favaro L, Gamba M, Gili C, Pessani D. 2017 Acoustic correlates of body size and individual identity in banded penguins. *PLoS ONE* **12(2)**, e0170001. (doi: 10.1371/journal. pone.0170001)

[26] Altmann J. 1974 Observational study of behavior: sampling methods. *Behaviour* **49**, 227–267.

[27] Boersma P. 2001 Praat, a system for doing phonetics by computer. *Glot Intern.* **5**(9/10), 341-345.

[28] Bates D, Mächler M, Bolker BM, Walker SC. 2014 Ime4: Linear Mixed-Effects models using Eigen and S4 R package. Version 11–7. (http://CRANR-projectorg/package=Ime4)

[29] R Development Core Team. 2008 *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org.

[30] Dobson AJ. 2002. *An introduction to Generalized Linear Models*. Boca Raton: CRC Press.

[31] Estienne V, Stephens C, Boesch C. 2017 Extraction of honey from underground bee nests by central African chimpanzees (*Pan troglodytes troglodytes*) in Loango National Park, Gabon: Techniques and individual differences. *Am. J. Primatol.* **79(8)**, e22672. (doi: 10.1002/ajp.22672)

[32] Barr DJ, Levy R, Scheepers C, Tily HJ. 2013 Random effects structure for confirmatory hypothesis testing: keep it maximal. *J. Mem. Lang.* **68**, 255–278. (doi: 10.1016/j.jml.2012.11.001)

[33] Ferrer-i-Cancho R. 2016 Compression and the origins of Zipf's law for word frequencies. *Complexity* **21**, 409–411. (doi: 10.1002/cplx.21820)

[34] Hřebíček L. 1995. *Text levels: Language constructs, constituents and the Menzerath-Altmann Law*. Verlag Trier: Wissenschaftlicher.

[35] Ferrer-i-Cancho R, Hernández-Fernández A. 2013 The failure of the law of brevity in two new world primates. Statistical caveats. *Glottotheory* **4**, 45–55. (doi: 10.1524/glot.2013.0004)

[36] Bowling DL, Fitch WT. 2015 Do animal communication systems have phonemes? *Trends Cogn. Sci.* **19**, 555–557. (doi: 10.1016/j.tics.2015.08.011)

[37] McComb K, Reby D, Bake L, Moss C, Sayialel S. 2003 Long-distance communication of acoustic cues to social identity in African elephants. *Anim. Behav.* **66**, 317–329. (doi: 10.1006/anbe.2003.2047)

[38] Gustison ML, Townsend SW. 2015 A survey of the context and structure of high- and low-amplitude calls in mammals. *Anim. Behav.* **105**, 281-288. (doi: 10.1016/j.anbehav.2015.04.021)

[39] Shannon CE, Weaver W. 1949 The Mathematical Theory of Communication. Urbana, IL: University of Illinois Press.

[40] Lengagne T, Aubin T, Lauga J, Jouventin P. 1999 How do king penguins (*Aptenodytes patagonicus*) apply the mathematical theory of information to communicate in windy situations? *Proc. R. Soc. Lond. B* **266**, 1623-1628. (doi: 10.1098/rspb.1999.0824)

[41] Favaro L, Pichegru L. 2018 Penguins: Behavioural ecology and vocal communication. In Encyclopedia of Animal Cognition and Behavior (eds J Vonk & TK Shackelford), pp 1-9. Springer International Publishing. (doi: 10.1002/ar.1091110305)

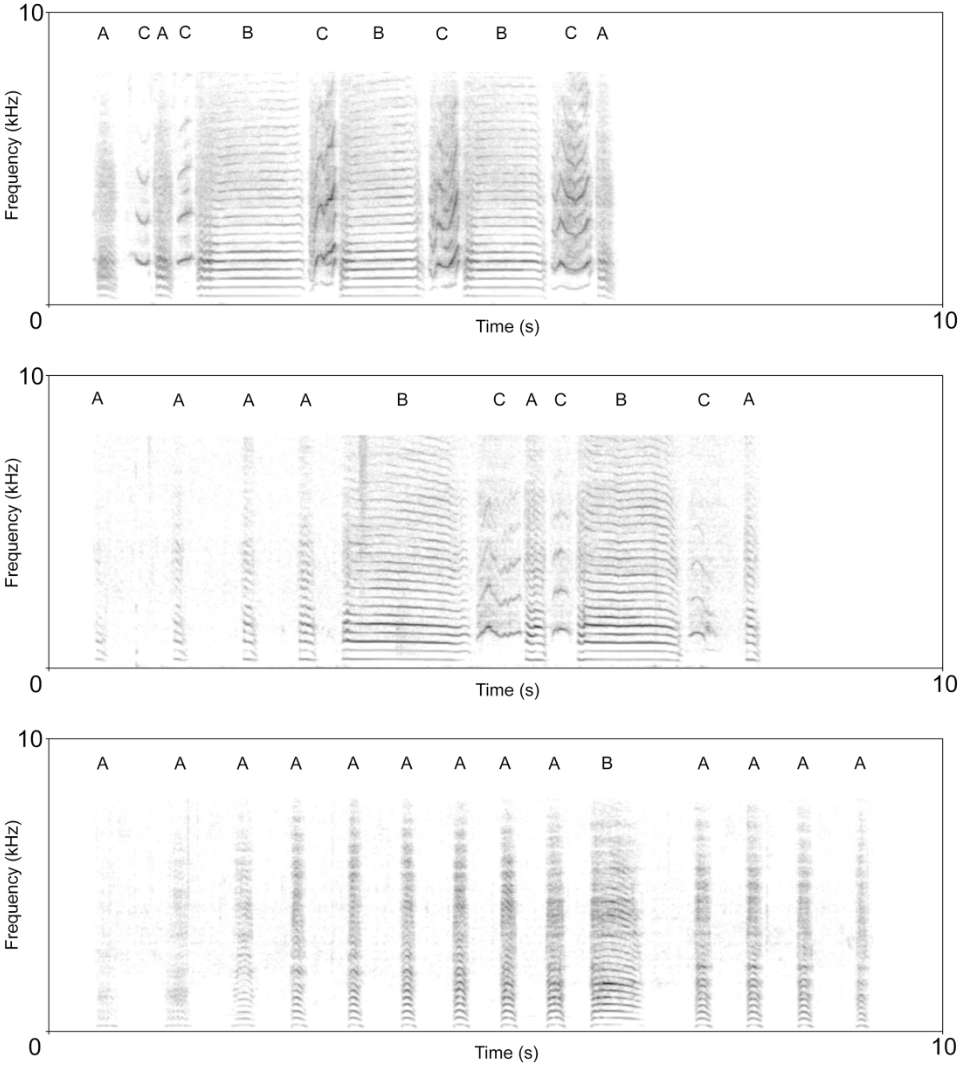

Figure 1. Example of variation in number and duration of syllables in ecstatic display songs of three adult African penguin males. Spectrograms were obtained in Praat using a Gaussian window shape, window length = 0.05 s, number of time steps = 1000, number of frequency steps = 500, dynamic range = 55 dB. Capital letters indicate the three syllables types.

119x139mm (600 x 600 DPI)

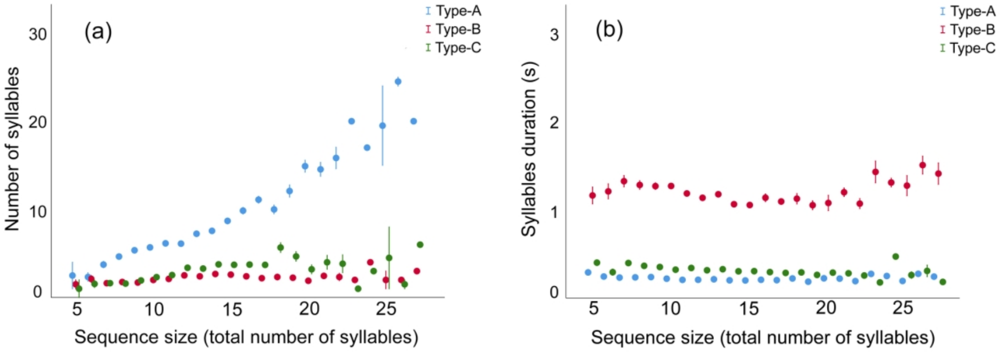

Figure 2. Relationship between the number (a) and the duration (b) of the syllables of each type and the sequence size. Points indicate mean values ± SE.

170x64mm (300 x 300 DPI)