Dolphin communication. A quantitative linguistics approach

Author:
Amalia Alkisti Vradi

Supervisor:
Ramon Ferrer-i-Cancho
Department:
Computer Science

Master in Innovation and Research in Informatics,
Data Science specialization

Facultat d'Informàtica de Barcelona (FIB)
at Universitat Politècnica de Catalunya (UPC) - BarcelonaTech

Date of defense:
April 28, 2021

Abstract

Comparative studies between human languages and animal communication have revealed shared statistical patterns that can shed light on the principles that govern communication across species while establishing the foundations to understand the evolution and the origin of languages. Two linguistic laws - Menzerath's law and Zipf's law of abbreviation - provide the framework to study the shared principle of information compression. Menzerath's law states that the longer the construct, the shorter its consistent parts, while Zipf's law posits a negative correlation between signal length and frequency of use. These statistical patterns are found in complex behaviours across diverse taxa, suggesting that the principle of compression is universal in animal communication. Here, we investigate whether the whistle of dolphins (*Tursiops truncatus*), a species widely known for its outstanding communication and social skills, conform with these linguistic laws. We show that, in dolphin vocal sequences, there is a negative relationship between the number and the duration of whistles, in line with Menzerath's law. Furthermore, based on an unsupervised whistle type classification, we find patterns that are consistent with Zipf's law of abbreviation in the relationship between the duration of a whistle type and its frequency of use. These findings provide evidence for coding efficiency in the vocal communication system of this species and for the first time among cetaceans. Finally, our results suggest that compression underpins human and dolphin vocal communication, illustrating the importance of recent extensions of information theory and also the need of exploring linguistic laws beyond human vocal systems.

Keywords: quantitative linguistics, Menzerath's law, Zipf' law of abbreviation, compression, animal communication, linguistic universals.

Contents

List of Figures

List of Tables

1 Introduction

Understanding the evolution of communication and the origin of languages has intrigued philosophers and scientists for centuries. Human languages have been considered unique, complex forms of communication that are characterised by syntax, semantics and the combination of these in various ways to produce different meanings. However, recent studies have shown that these characteristics are not uniquely human. For example, it has been shown that birds can produce calls with specific meanings and combine them into sequences that exhibit compositional syntax [Suzuki et al., 2020]. Evidence like this has directed the research interest on the evolution of human communication in comparing quantitatively human languages with other non-human communication systems. The framework for performing such comparative studies has been offered by quantitative linguistics.

A central goal of quantitative linguistics is to identify fundamental principles that are shared across all languages. A cornerstone in this effort is the investigation of linguistic laws, i.e. the statistical regularities that emerge in human languages. Among them, the most prominent ones are Menzerath's law and Zipf's law of abbreviation (law of brevity in short). The quantitative linguistic observations that these laws underline have been linked to compression [Ferrer-i Cancho, 2016].

In information theory, the principle of compression is expressed as the problem of minimising the expected length of a code [Ferrer-i-Cancho et al., 2013]. In that framework, this proposition has been formulated as following

$$L = \sum_{i=1}^{V} p_i l_i \tag{1}$$

where p_i and l_i are respectively the probability and the length of the ith element of a repertoire of size V [Cover and Thomas, 2006]. Solving this cost function for minimizing L consists of finding the lengths given that the probabilities of the elements are not changing and that all codes are unique [Gustison et al., 2016]. These constraints indicate a relationship between probability and length that cannot be positive [Ferrer-i Cancho et al., 2020]. Therefore, the minimization of L predicts a negative correlation between frequency and length, which is the relationship indicated by Zipf's law of abbreviation. [Ferrer-i-Cancho et al., 2013]. The same arguments here have been extended in order to be applied to Menzerath's law by generalizing function 1 [Gustison et al., 2016].

In the last decade, through the study of these laws, there have been established many parallels between human languages and animal communication. Menzerath's law manifests in language as a tendency of the mean size of the parts to decrease as the number of parts increases [Altmann, 1980]. Evidence of this statistical pattern has been reported in the communication system of diverse animal taxa, such as in gibbon songs [Clink et al., 2020], gelada's vocal sequences [Gustison et al., 2016] and chimpanzee gestures [Heesen et al., 2019]. Equivalently, Zipf's law of abbreviation, as commented earlier, dictates a negative relationship between frequency of use and duration, which is translated as a tendency of the most frequent types to be shorter. The statistical patterns of this law had been reported for the first time in the calls of chickadees [Hailman et al., 1985] and subsequently in other communication systems and behaviours, such as in Formosan macaques [Semple et al., 2010], again in the gestural repertoire of chimpanzees and in the surface behaviours of dolphins [Ferrer-i-Cancho and Lusseau, 2009].

Dolphins specifically are a particularly interesting species to study from the family of mammalians. They have attracted the interest of researchers for decades, providing a plethora of studies that shed light on their physical characteristics, cognition abilities and social structure. Today we know that dolphin brains are among the larger ones in non-human animals. Their brain to body ratio, which is considered a physical measure of intelligence, is classified as second after humans [Marino et al., 2007]. Another measure of intelligence, that of tool usage, is also documented in dolphins, as they use sponges while hunting for their prey and appear to transmit this technique culturally [Krützen et al., 2005]. Dolphins show many additional striking similarities to humans. They have shown capabilities of vocal learning and vocal mimicry [Reiss and McCowan, 1993], and have shown as well evidence of mirror self-recognition [Reiss and Marino, 2001]. Despite though their thorough study over the years, the communication system of dolphins still puzzles researchers.

The main objective of this study is to investigate the communication system of dolphins with the aim to identify parallels with human communication. We perform quantitative linguistic analyses on this species as a means of finding indicators of potential structure in their communication system [McCowan et al., 2005]. Initially, we test for evidence that dolphin whistle sequences adhere to Menzerath's law. For conducting these experiments, it was necessary to address the challenges of segmenting the vocalisations into sequences. The typical approach of studies that are dealing with this kind of task has been to designate a minimum duration threshold as a break point between sequences (e.g. [Clink et al., 2020, Gustison et al., 2016]). Here, we propose a new methodology that instead of one fixed threshold, it considers a range of values that are meaningful to the species in study. We find that in dolphin sequences, the longer a phrase (in terms of the number of consisting whistle signals), the shorter the individual whistles would be.

Next, for the study of Zipf's law of abbreviation, the analysis focused on individual dolphin whistles, where we dealt with the problem of clustering them into types. We replicated an unsupervised clustering methodology to separate the whistles into different repertoires, and we ran our experiments across a range of different configurations. In accordance with the law of brevity, we found that the frequency of use of a type has a negative relationship with the whistle duration. Finally, through the experiments of both laws, we address the secondary objective of this study, which is to investigate the relationships predicted by the theory of compression.

These analyses provide the first evidence of these two linguistic laws in the signal communication system in the family of cetaceans. The work that is presented in the following chapters is an additional effort in the continued exploration of the applicability of statistical laws in non-human systems. The evidence that we find here in a system so distinctive from ours could provide insights regarding the evolution of universal linguistic patterns and the diversity we see today in languages.

The remainder of this thesis is divided into five chapters. Chapter 2 presents the necessary information for understanding the nature of the dolphin whistles data and the data sets we are using, as well as some interesting preliminary results regarding the parallels of human and dolphin communication. In chapter 3, we include our analysis on Menzerath's law, our novel methodology in separating dolphin whistles into sequences and the results we obtain when we test for the relationship between whistle duration and sequence size. In chapter 4, we revisit the task of clustering dolphin whistles into types where we replicate a two-phase clustering methodology developed to address this challenge. Chapter 5 shows the results we obtained in our research on the relationship between the frequency of whistle types and their duration. Finally, chapter 6 discusses the implications of our findings for the debate of linguistic universals and our understanding of communication.

For the organization of Chapter 3, 4 and 5 we borrow the style of top general science journals such as PNAS[1] and Nature Communications[2]. According to this format, our findings are represented in the order of Introduction, Results, Discussion and lastly Methods.

---

[1] `www.pnas.org`

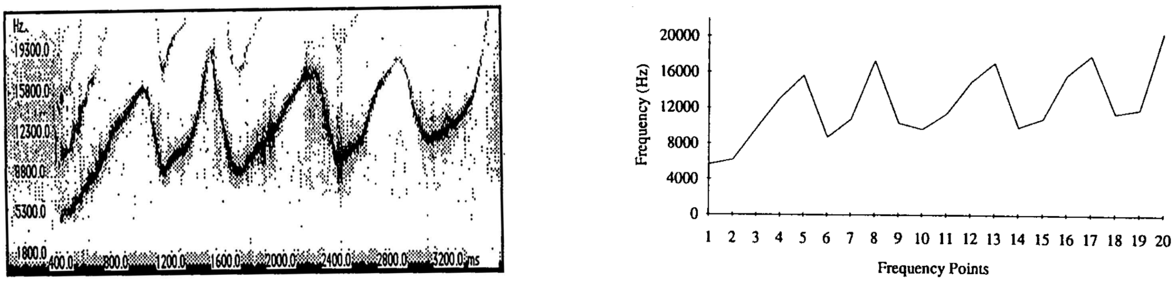
[2] `www.nature.com/ncomms/`

2 | Dolphin whistles

Dolphin vocalisations are generally categorised into two types of sound emissions, pulsed and tonal sounds [Caldwell and Caldwell, 1968]. The most prominent type of the first category, pulsed sounds, is the click, a short-broadband sound [Au, 1993] that has been associated mainly with echolocation [Freitag and Tyack, 1993]. However, it is the other category of dolphin vocalisation that has been associated with communication. Narrow-band, tonal sounds that are generally referred to as whistles.

Dolphin whistles are typically presented and explored with time–frequency spectrograms. Spectrograms provide a visual representation of the long and short term characteristics of the vocalisations and capture their distinct features. For the case of dolphin whistles, spectrograms are particularly useful as they can highlight the vocalisation's relative change in frequency over time, which is known as the 'whistle contour'. The whistle contour then can be used to study in depth the dolphins' vocal repertoire and their communication system.

Whistle contours are represented as time series. Compared to other methods that use single measurements as features to characterise the whistle, such as the mean frequency or the duration, for the whistle contour we extract a series of fundamental frequency points of the vocalisation. For example, in figure 1b, we see the equivalent whistle contour as a time series of the vocalisation from figure 1a. In this example, the time series is consisted of 20 frequency points.



(a)

(b)

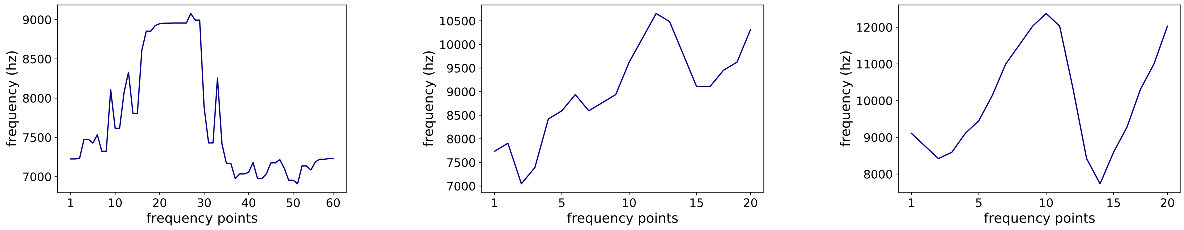
Figure 1: Figure *(a)* shows an example of a whistle spectrogram. In *(b)* we see equivalent whistle contour as time series. Figure taken from McCowan et. al. [McCowan, 1995]

In our work, we incorporated two data sets of dolphin whistles which have been encoded as whistle contour time series. The first is a very long whistle catalogue, which contains more than 160k dolphin whistles. The second one is a smaller data set that contains whistles from different age groups, adult and infant dolphins. Below, we briefly provide their summary information.

2.1 Data acquisition

We analysed a data set of 163.436 whistles that were collected by Brenda McCowan over a period of 5 months during the year 1997. The whistles were produced by four individual dolphins in captivity which were recorded by a hydrophone 24 hours a day. The collected recordings were then transformed into spectrograms and segmented into separate whistles based on the silent gaps between the vocalisations. Finally, across the duration of each whistle, evenly separated frequency points were extracted to produce the equivalent whistle contour time series, as described earlier. The time series here contain 60 frequency values for every whistle. The final content of this data set included the whistle contours as time series organised into 187 files, each of which is distinguished with a recording id.

This automation design of the process allowed the researchers to build a data set of dolphin whistles of such magnitude with the technologies available at that time. Due to this automatic process of collecting the recordings though, the data set does not contain information about the dolphin identity for each vocalisation, as typically, in animal studies the identity of the vocaliser is assigned to the data by an observer during the period of the recordings.



(a) (b) (c)

Figure 2: Whistle contour examples. In figure *(a)* we see a contour example that belongs to the big collection of dolphin whistles. The other two are examples of the additional data we used, where the whistle contour in *(b)* corresponds to an adult dolphin and (c) to an infant.

2.2 Whistle duration

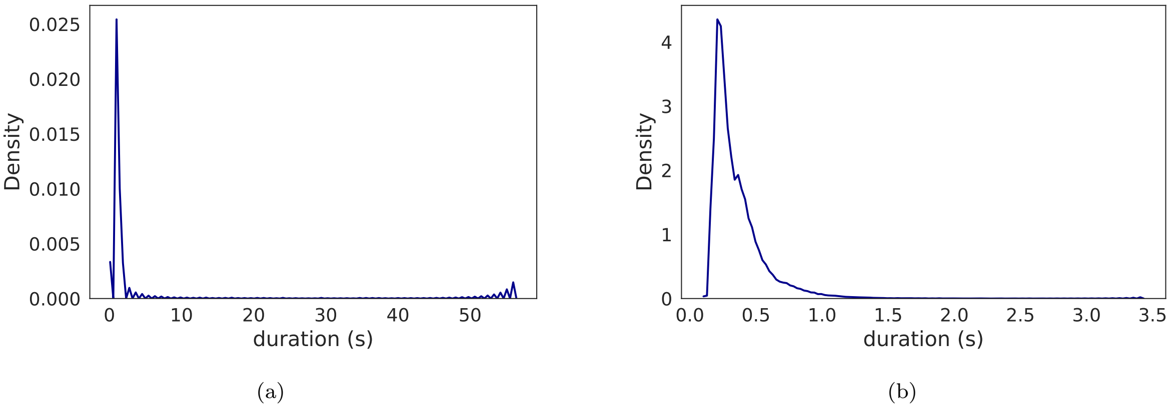
As dolphins have been extensively studied over the years, there are plenty of previous works that provide information about the observed characteristics of dolphin whistles. Particularly for the whistle duration, previous research studies agree on their findings. In one of their first works, Cladwell & Cladwell [Caldwell and Caldwell, 1968] reported that the average duration of five specific whistles ”signatures” were in the range of $[0.79, 0.87]$ seconds. Today, in more recent works, Buckstaff [Buckstaff, 2004] reports that whistles have an overall duration of 0.01 to 4 seconds. Additional work from Petrella [Petrella et al., 2012] on common dolphins *(Delphinus sp.)* reports similar findings on the duration of whistles, which range from 0.01 to 4.00 s (mean±SD: 0.27±0.32).

	N	*Mean*	*Median*	*St.Dev*	*Min*	*Max*
initial	163343	0.36	0.29	0.54	0.15	53.38
final	163268	0.35	0.29	0.20	0.15	3.39

Table 1: Summary statistics of whistles duration (in seconds) of the initial data set of 163343 dolphin whistles. The table contains the summary before removing problematic data entries from the data set (initial) and final statistics after cleaning the data (final).

During our exploration of the data set, we found evidence of data that were not in accordance with the studies mentioned earlier. The general summary statistics for the whistle durations were a bit off in the case of the maximum durations. Table 1 contains these summary statistics as they were initially, and figure 3a shows the initial density distribution plot for the duration.

After analysing further the reported data with the maximum duration, we found that only the last 60 whistles out of the total of 163343 (when ordered according to duration) had values in the range of [4.8, 56.37] seconds, showing that only a very small number of data accounted for the extremely big values we reported above. Exploring further the data and plots of the whistles in the range between 3.6 to 4.8 seconds we still saw shapes that did not resemble whistle contours. Example cases of these particular entries are shown in appendix B in figure 15. About the shape of the contour, Markov et al. [Markov and Ostrovskaya, 1990] in their analysis of dolphin signals report that in a typical whistle we see on average 5-7 blocks, which in some cases can reach 12. Here, with blocks they are referring to the different phonations or peaks in a whistle.



(a)

(b)

Figure 3: Density plots of whistle duration. In figure *(a)* depicts all the initial data. In *(b)* we see the distribution after discarding the problematic entries.

Summarizing our observations above, we concluded that these cases did not constitute proper whistles, resulting in the exclusion of 75 whistles from our data set. The effect of this did not have a significant change in the reported statistics, as we confirm with the summary table 1. However, it did have an important effect on the plots of the duration, as after their removal, we obtain a clear picture of the distribution of whistles duration.

Figure 3 contains the density plots of whistle duration before and after the removal of the problematic entries. We see that duration values are concentrated around 0.3 seconds. Additionally, the mass of the distribution is concentrated on the left of the figure and it has a long tail on the right, giving a right-skewed distribution. In the plot 3b, we observe two peaks, the one mentioned above at 0.3 and a second one at 0.5 seconds.

It is particularly interesting to mention that the same findings are reported by Markov and Ostrovskaya [Markov and Ostrovskaya, 1990] in their analysis of dolphin signals (figure 7 in [Markov and Ostrovskaya, 1990]). In their work, they comment that the analysis of the dolphin whistles distribution always has two peaks, each of which has a different origin. They mention that the first peak is associated with "informational interaction of animals", while the second one is related to a stressed state of the vocalizer.

Additional interesting results we observe in figure 4a, where we plot the density of the whistles duration using a logarithmic scale on the y-axis. The figure shows an exponential decay in the durations, a behaviour which is also found in human languages [Corral and Serra, 2020].

While this statistical behaviour in humans had been traditionally studied using written communication, recent studies show that the shame principles hold for the physical units in oral communication [Torre et al., 2019, Torre et al., 2017]. This finding provides an important base for the comparative studies between animal and human communication systems, as we can compare units using similar physical characteristics, such as the duration of a signal.

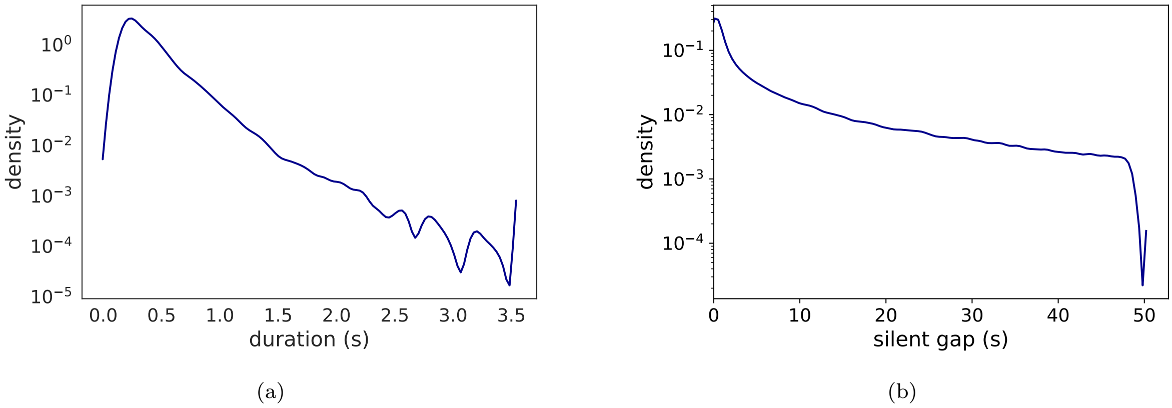


Figure 4: Distribution of whistle duration and silences. Figure *(a)* shows the density distribution of whistle duration when we use log scale in y-axis. We used a smoothing factor h=0.05 for producing this plot to soften the effect of the extreme values at the right end of the tail. In figure *(b)*, we plot the density distribution of silences between the whistles.

Silent gap between whistles

An additional duration reported in the initial data set is that of the silent gap between two consecutive whistles. The reporting of these silence intervals is important for our study, as this information is later used for segmenting long series of whistles into a series of sequences of whistles. In figure 4b, we plot the density distribution of the silences and in table 2 we report their summary statistics. During the data acquisition process, the duration of silences was calculated in milliseconds up to a certain threshold that was probably determined by numeric limits of the computer architecture used at that time. If a silence was passing that threshold, the value was replaced with a negative one and was considered to be more than a few minutes. For this reason, we had available the specific value for the silent interval for 153721 cases out of 163342. For the remaining cases, we considered them to have a duration longer than a minute, and we did not include them in the calculation of the summary statistics.

	N	*Mean*	*Median*	*St.Dev*	*Min*	*Max*
silent gap	153721	6.16	1.38	10.05	0.0	48.69

Table 2: Summary statistics of silences (in seconds) between the whistles. The statistics were calculated using N values, as for the rest, the silence was considered to be more than a minute and their exact values were not available.

2.3 Adult and Infant dolphins data

We used an additional data set that contains whistles of adult and infant dolphins. The data sets includes the identity of dolphins and their sex, the whistle type, the whistle duration and its frequency points, as well as the behavioral context in which the whistle was recorded. Finally, it also indicates the sequence in which the whistles were part of. Detailed methods for recording, analysing and categorising the dolphin whistles can be found in the works of McCowan and Reiss [McCowan, 1995, McCowan and Reiss, 1995a, McCowan and Reiss, 1995b]. In figure 2b and 2c, we see example whistles from the adult and infant dolphin data respectively.

	N	*Mean*	*St.Dev*	*Min*	*Max*
adults	434	0.37	0.34	0.03	3.74
infants	964	0.39	0.29	0.06	2.95

Table 3: Summary statistics of whistle duration for adult and infant data sets

The data set of adults includes individually identified whistles from three males (Bay, Gor, Sch) and five females (Che, Cir, Sad, Sto, Ter) with a total of 434 whistles, 27 whistle types and 130 sequences. In the case of infants, the data set includes the individually identified whistles from eight males and one female (Tas) with a total of 964 whistles, 103 whistle types and 289 sequences.

A summary of the elementary statistical properties of the sequences is provided in Table 4. The sequences include whistles that appeared at least once, and both data sets do not include sequences with length smaller than 2. Finally, for the case of whistle duration, there were 28 missing values in the adults data set and 140 in the data of infants.

Adult				Infant			
Name	*Whistles*	*Types*	*Seq*	*Name*	*Whistles*	*Types*	*Seq*
BAY	27	10	8	DEL	65	18	24
CHE	110	12	34	DES	40	3	7
CIR	47	4	15	ECB	15	5	6
GOR	6	6	2	LIB	334	31	102
SAD	110	13	33	NEP	53	9	9
SCH	19	4	6	NOR	235	39	74
STO	94	12	27	PAN	148	42	43
TER	21	7	5	SAM	39	4	12
				TAS	35	9	12

Table 4: Summary of the elementary statistics of sequences for each dolphin. The table summarizes the total number of whistles per dolphin, the number of different whistles types and the total number of sequences.

3 | Menzerath-Altmann law

Menzerath–Altmann law, shortly Menzerath's law, is a linguistic law that states that "the larger the construct, the smaller the size of its constituents" [Menzerath, 1954, Altmann, 1980]. In its original formulation, this law has been used to describe the statistical patterns found in written communication. Evidence has shown that the law holds in different scales of analysis. For example, the longer a sentence, the shorter its consisting words tend to be. Similarly, the longer a word, the shorter its consistent syllables. Beyond linguistics, Menzerath's law has been applied in a wide range of systems where it confirms again the negative correlation between constructs and constituents, such as in music [Boroda and Altmann, 1991], in genomes at different levels of organization [Ferrer-i-Cancho and Forns, 2010, Hernández-Fernández et al., 2011] and in proteins [Shahzad et al., 2015].

Beyond human communication systems, patterns consistent with this law have been reported in a few other species: in gelada vocal sequences [Gustison et al., 2016], in male gibbons [Clink et al., 2020], penguin songs [Favaro et al., 2020], and chimpanzee hoots [Fedurek et al., 2017] and gestures [Heesen et al., 2019]. The law has been linked several times with the principle of compression in the information-theoretic sense and in the work of the communication system of geladas the authors offer formal support for this link, providing a mathematical formulation [Gustison et al., 2016]. The evidence so far has led to the consideration that compression is a universal principle in communication. However, it is repeatedly stressed that further testing for adherence to the linguistic laws in more systems is mandatory to facilitate the exploration of universal properties of communication [Gustison et al., 2016, Heesen et al., 2019, Clink et al., 2020].

In this analysis, we test for Menzerath's law in dolphin whistle sequences. Dolphin whistles here are not used as an equivalent to human words but rather as a recognized unit in the communication system of dolphins. For delimiting the boundary between consecutive sequences, we used the information about the silence intervals between the whistles. Specifically, we defined an interwhistle interval δ as the maximum silent gap between two whistles that belong to the same sequence. If the silence between the whistles is lower than δ, then we assign them to the same sequence.

A priori, we do not have knowledge about the "true" whistle sequences, as this concept is not well defined in non-human communications. In fact, the debate about the definition and classification of acoustic sequences is still open [Kershenbaum et al., 2013]. An approach in constructing whistle sequences has been to define a fixed interwhistle or intersequence interval and build the sequences using them as a reference [McCowan et al., 1999, Gustison et al., 2016]. In this analysis, we want to study the effect of this definition in the results we obtain. Therefore, we built different sets of sequences for a range of values of δ. We then created three additional randomized versions of these initial sets of sequences by shuffling once the whistle durations, once the silent gaps and once both (see chapter 3.3 for further details on the randomization procedure).

Firstly, we tested for a correlation between the number of whistles within a sequence and the mean duration of these whistles, where a negative relationship would provide evidence for Menzerath's law [Gustison et al., 2016]. In continuation, to assess in detail the effect of sequence size on the whistle duration, we fitted a series of generalized linear models where we measured the importance of the sequence length and the whistle's position in the sequence.

Shuffling methods

Initially, we built sets of sequences, choosing for each set a different interwhistle interval within the sequence δ in the range of 0.1 to 10 seconds, for every 0.1 seconds. Elementary statistics for the sequences lengths for all 100 sets produced with different intervals can be found in table 14 in appendix B.

We then repeated the same process three times to construct sets of random sequences. Firstly, we maintain the distribution of individual whistles, their order and their durations but we shuffle the silence between the whistles. As before, we used 100 values for the δ and obtained new sets of random sequences. We performed the same steps after shuffling the durations of the whistles but maintaining the silences order between them. Finally, we constructed 100 more sets of sequences for which we shuffled both the silence between the whistles as well as their duration. Details about the total number of sequences obtained for every set and shuffling, as well as elementary statistics, are included in the appendix B.

The first two shuffling scenarios are maintaining some aspect of the initial data (once the order of silences and once the order durations), while the last one shuffles all parts. All three shuffled sets of sequences are used in the following analyses as a comparative tool, to control for arbitrary structure in the data that might result in patterns that are consistent with the law. Furthermore, we introduce for the first time in this work the methodology of shuffled sequences to study the argument of the inevitability of linguistic laws.

For the sets of sequences constructed after shuffling the silences between the whistles, we analyzed the distribution of the length of the sequences across all the values of δ used. In equation 2, the expected value $E(Y)$ was used to infer the value of p, where Y is the mean sequence length (mean number of whistles in sequences) and p is the sequences' probability to have a certain mean length. Subsequently, the expected variance was calculated (eq. 3). Figure 5 shows that the line corresponding to the values for the shuffled silence sets coincides exactly with the line of the geometric distribution.

$$E(Y) = \frac{1-p}{p} \tag{2}$$

$$var(Y) = \frac{1-p}{p^2} \tag{3}$$

This result suggests that in the process of generating sequences in the case of shuffled silences, the size of the produced sequence does not have any effect on the probability of ending the sequence, a ”memoryless” process [Degroot, 1986]. While the fact that we do not see this result in the initial set of sequences where no shuffling was performed, provides evidence that there is in fact memory of length in the process of generating sequences.

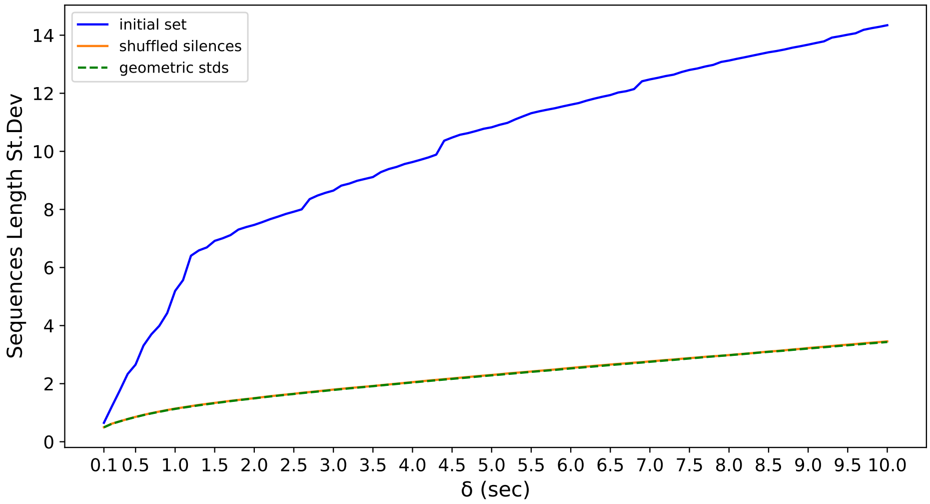


Figure 5: Standard deviation of sequence length across all deltas for initial sequence sets and the sets after shuffling the silences between the whistles. With dashed green is the geometric distribution where the parameter has been obtained by calculating the expected value using the formula of the expected variance (Equations 2 & 3)

3.1 Results

Whistle duration and sequence size correlation

To test for adherence of Menzerath's law we firstly ran correlation tests between the mean whistle duration per sequence and the sequence length (how many whistles is the sequence consisted of). In figure 6, we plot the Spearman correlation results for all the sets of sequences, for all values of δ. Results that are consistent with the law are those that return a negative relationship. We found such negative correlations for the sequences that were constructed with values of δ in the range of 0.1 to 0.3 seconds. These results were obtained only in the scenario where we did not perform any shufflings. For all the other scenarios and values of δ, the test returned positive correlations. Finally, all the experiments produced significant p-values below the threshold of 0.001.

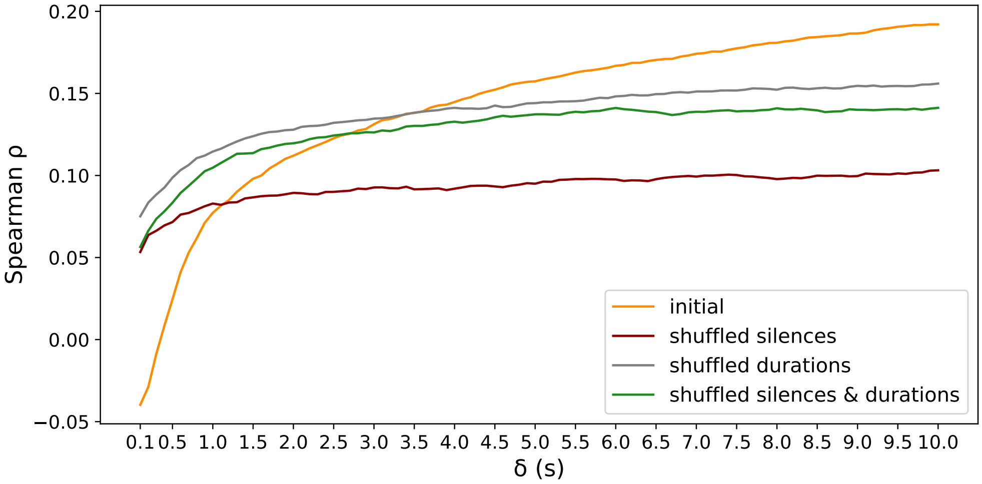


Figure 6: Spearman correlation results for all sets of sequences. Negative correlation, consistent to Menzerath's law, is observed for values in [0.1,0.3] seconds for the initial set of sequences (δ=0.1: ρ=-0.039, p<0.001, δ=0.2: ρ=-0.028, p<0.001, δ=0.3: ρ=-0.0085, p<0.001). The sequences constructed after shuffling the silences, duration of whistles or both do not show negative correlation for any of the δ

To exclude the case of the findings being an artefact of trivial scaling, we ran a complementary analysis testing the relationship between the sequence size and the total duration of its consisting whistles [Ferrer-i Cancho et al., 2014]. We obtained a significant positive correlation, which excludes the trivial case of the adherence to the linguistic law [Heesen et al., 2019].

LMM statistical analysis

In continuation, we tested for the existence of the law by fitting linear mixed model analyses and based our evaluation on the Akaike information criterion. We fit models for all values of δ and for all the sets of sequences that we constructed. For each δ, the full model contained the individual whistle duration as the response variable, and the sequence length and whistle's position in the sequence as fixed effects. We included the recording id as a random effect. For the null model, we included only the random effect without any fixed effect. Patterns consistent with the law are those that have a negative coefficient for the estimate of sequence length.

For the initial sets of sequences, we obtained the sequence length as a negative predictor of the whistle duration for the case of δ=0.1 seconds, which gives support for adherence to Menzerath's law. Additionally, this negative coefficient was confirmed both by the estimate's standard error as well as by the calculation of the confidence intervals. Table 5 contains in detail this model's output, and figure 7a shows the plot of both estimates of the model along with their confidence intervals. This result is in accordance with the correlation analysis we performed previously for the sequences of the same δ.

In contrary, for all the remaining values of δ, the models returned a positive coefficient for the sequence length, suggesting a positive relationship with the whistle duration. We summarize the model outputs for the values of δ in tables that can be found in the appendix C. Figure 7c shows some of the model results we obtained for various values of δ, where the results indicated a positive relation between the sequence length and the whistle duration.

		Confidence intervals	
Predictors	Estimate ± SE	Lower	Upper
Intercept	342.01 ± 5.02	332.11	351.86
Sequence length	−2.02 ± 0.55	−3.09	-0.95
Position	5.06 ± 0.84	3.41	6.70
Observations	163,268		
AICd	34.47		

Table 5: LMM table results for $\delta = 0.1(s)$. The sequence length estimate (± confidence intervals) shows a negative relation between this predictor and whistle duration. AICd indicates the difference between the the AIC of the full and null model (AIC null - AIC full).

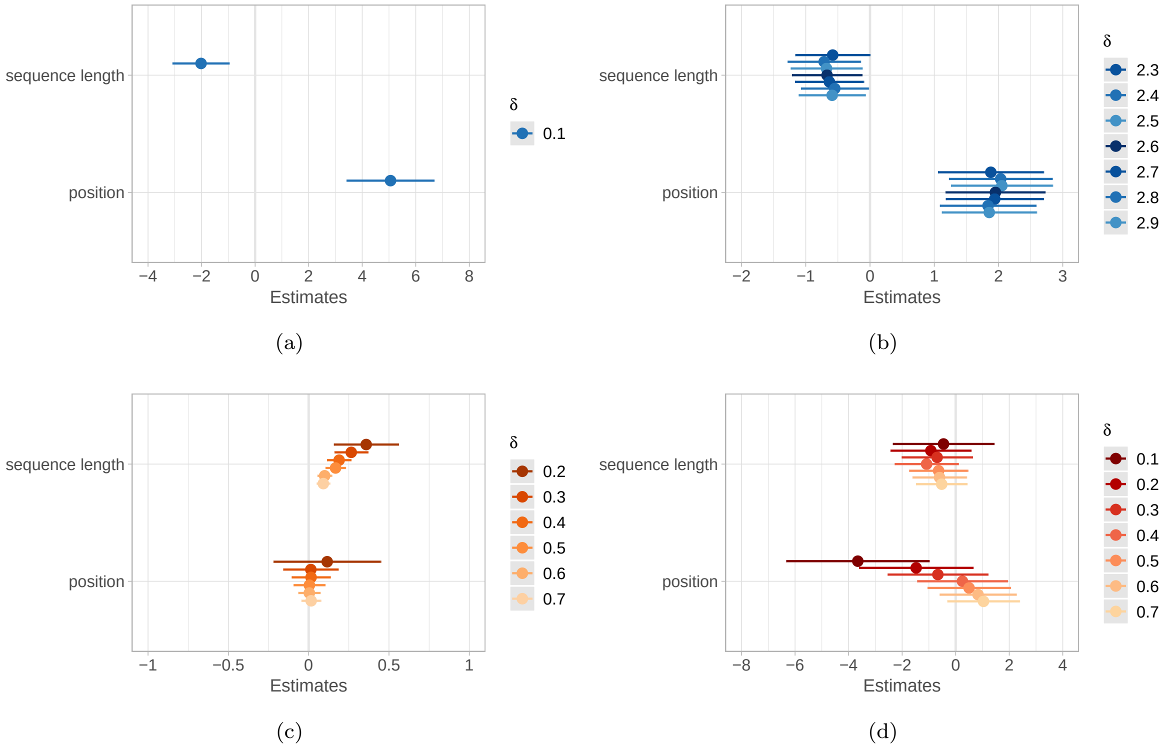
			Confidence intervals		
δ	Predictors	Estimate ± SE	Lower	Upper	AICd
2.3	Intercept	339.94 ± 5.04	330.03	349.82	18.95
	Sequence Length	-0.58 ± 0.30	-1.16	-0.00	
	Position	1.88 ± 0.42	1.06	2.70	
2.4	Intercept	340.19 ± 5.03	330.27	350.07	22.66
	Sequence Length	-0.71 ± 0.29	-1.28	-0.15	
	Position	2.03 ± 0.41	1.23	2.84	
2.5	Intercept	340.02 ± 5.03	330.10	349.90	24.87
	Sequence Length	-0.68 ± 0.28	-1.23	-0.12	
	Position	2.05 ± 0.40	1.26	2.84	
2.6	Intercept	340.10 ± 5.03	330.19	349.98	22.73
	Sequence Length	-0.67 ± 0.28	-1.21	-0.12	
	Position	1.95 ± 0.40	1.18	2.73	
2.7	Intercept	339.97 ± 5.03	330.05	349.84	24.05
	Sequence Length	-0.63 ± 0.27	-1.17	-0.10	
	Position	1.94 ± 0.39	1.18	2.70	
2.8	Intercept	339.78 ± 5.03	329.87	349.65	22.86
	Sequence Length	-0.55 ± 0.27	-1.07	-0.02	
	Position	1.84 ± 0.38	1.09	2.59	
2.9	Intercept	339.89 ± 5.03	329.97	349.76	23.56
	Sequence Length	-0.59 ± 0.26	-1.11	-0.07	
	Position	1.86 ± 0.38	1.12	2.59	

Table 6: Results of linear mixed-model analysis sequences that were constructed after shuffling the order of silences between the whistles and their durations.

For the sets of sequences which were constructed after performing some of the shuffling of the whistles information, we obtain a different picture. When shuffling silences, for almost all the values of δ, the estimate of sequence length was positive. The models for values of δ in

$[0.1, 0.9]$ were the only ones that returned a negative value for the estimate of sequence length. However the performance of these models according to AIC was worse than the null model, indicating that none of these results were reliable. Similar were the results for the case where we shuffled the durations of the whistles. All models were outperformed by the null model, which scored better AIC for all intervals, suggesting that we can not draw a conclusion about the existence of the law using any of them. Detailed summaries of all these models are included in the appendix C as well.

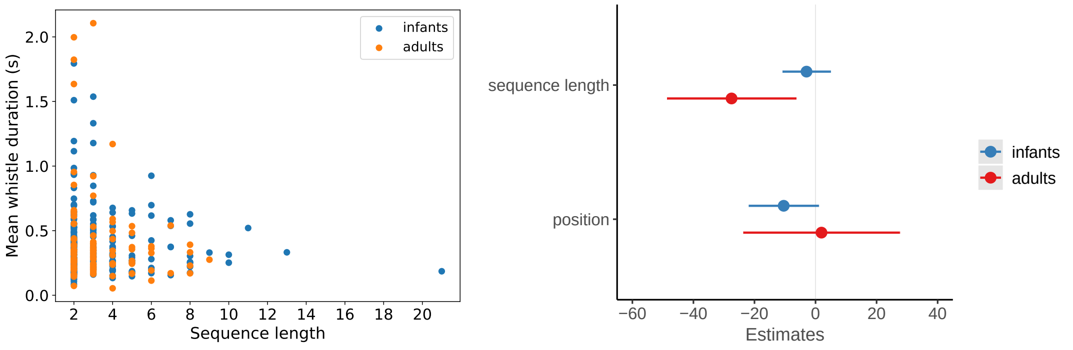
When shuffling both the silences and the durations of the whistles, we obtain different results depending on the interval that we are analyzing. All the coefficient estimates for the sequence length returned negative values, which however were not supported by the calculation of the confidence intervals, neither by the standard error. In figure 7d, we present examples of such model results in which we can see that the coefficient estimates have wide confidence intervals that suggest positive values. Finally, seven models for values of δ in the range of $[2.3, 2.9]$ returned a negative sequence length coefficient. These models were ranked higher from the null model according to their AIC and their confidence intervals agree on the negative result of the models. In table 6, we provide the detailed results of these models and in figure 7b we plot their coefficient estimates with their confidence intervals.



(a) (b)

(c) (d)

Figure 7: Figure *(a)*: coefficient estimates for δ=0.1 seconds from the set of sequences where no shufflings were made. Figure *(b)*: model results from the sets of sequences where we shuffled both the whistle duration and the silences between them. These models return results that are consistent with Menzerath's law. Figure *(c)* and *(d)* show example model outputs for cases where no patterns consistent with the law are found (i.e. the sequence length estimate is positive). The models in c belong the initial set of sequences, while in d to the set where we shuffled both durations and silences.



(a)

(b)

Figure 8: Figure *(a)*: scatter plot of sequence length against mean within sequence whistle duration in seconds for adult and infant dolphins. Figure *(b)*: plot of linear mixed-effects model coefficients for adult and infant dolphins.

Additional analysis: Adult and Infant dolphins dataset

Similarly as before, we performed the same analyses for the data of adult and infant dolphin whistles. The Spearman correlation tests showed a negative correlation between the sequence size and the mean whistle duration both for adults $(\rho = -0.101, p - value = 0.26)$ and infants $(rho = -0.035, p - value = 0.57)$, however without a significant p-value for any of them. In figure 8a, we plot the relationship between the sequence size and whistle duration.

For the linear mixed-model analysis, whistle duration was the response variable as before, and sequence length and whistle position were used as fixed effects. We used the id of the dolphins and their sex as random effects. The null model had the same random effects but no fixed effects. The model we fitted for the adult dolphins returned a negative coefficient for sequence length, providing further evidence for the existence of the law. In the case of infant dolphins, the model returned again a negative estimate for sequence length. However, its standard error and confidence intervals give a mixed result, both showing that the coefficient can take positive values. For all these models, we obtain wide confidence intervals which are related with the significantly smaller size of the data sets compared to the previous analysis. Figure 8 includes the plot of the coefficients of the models, and table 9 contains in detail the estimates output of the models.

			Confidence intervals		
	Predictors	Estimate ± SE	Lower	Upper	AICd
adults	Intercept	514.99 ± 67.61	381.16	654.14	5.52
	Sequence length	-27.51 ± 10.77	-48.59	-6.36	
	Position	1.949 ± 13.06	-23.65	27.55	
infants	Intercept	408.29 ± 31.93	343.35	472.06	7.83
	Sequence length	-2.95 ± 4.00	-10.81	4.88	
	Position	-10.43 ± 5.83	-21.85	0.99	

Table 7: Results of LMM used to test for relationships between whistle duration, sequence size and interval position for the data of adult and infant dolphins. AICd refers to the difference of the Akaike criterion between the null model and the full.

3.2 Discussion

We found that the dolphin whistle sequences follow statistical patterns consistent with Menzerath's law in multiple levels of our analyses. To begin with, we found a strong negative correlation between the number of whistles in a phrase and the mean duration of the whistles in the case of sequences constructed using δ=0.1. The initial negative relationship was obtained by the Spearman correlation tests and was confirmed by our analysis with the generalized linear models. This, however, was not the case for δ=0.2 and δ=0.3, for which only the correlation tests showed supporting evidence.

Furthermore, in our analysis with mixed-effects models, we obtained results consistent with Menzerath's law in the last set of the sequences we built, where we had shuffled both the whistles duration order as well as the silence order between them. However, the values of δ for which we obtain these results did not show any supporting evidence in any of the previous sets (initial, shuffled silences or durations) nor in the Spearman correlation tests, as was the case of $\delta = 0.1$. As these negative relationships appear in the randomized sets where we tried to shuffle all the building blocks of the sequences, we are skeptical about these results, and we consider the possibility that we actually created an artificial scenario that provided supporting evidence for the law.

In the way that we performed our analyses, we used the randomized versions of the initial sequences as a comparative tool to assess whether the existence of patterns consistent with the law is perhaps inevitable. If, for example, we obtained a negative relation between construct and constituents for δ=0.1 across all randomized cases, then we would not have been able to draw conclusions. Evaluating our results, the absence of supporting evidence in the randomized sets excludes the case that the presence of the law is inevitable. Additionally, the significant correlations that we found in our experiments earlier refute the view that the law is a result of trivial scaling [Solé, 2010].

Concerning the mechanism we developed for constructing the sequences, there are several points that are worth discussing. As mentioned earlier in the data chapter, the initial data set we used for the analysis does not contain the information of the dolphins' identity. The lack of this information can result in assigning in the same sequence whistles produced by different dolphins. The longer the time we allow between whistles when considering them as part of the same sequence (i.e. bigger values of δ), the greater the chances of including in the same sequence a whistle from another dolphin. On the contrary, with a shorter time value for δ, we reduce the chances of having multiple vocalizers in the same phrase. Therefore, although we can report evidence for adherence to Menzerath's law, we can not draw a conclusion about whether this accounts for individual dolphin phrases or for sequences on a population level. Considering again the results we obtain for δ=0.1, the smallest value tested for δ, an interesting question would be whether this small value for the silent gap manages to separate the sequences better than the rest of the values, without mixing whistles from different dolphins. At the same time though, it is possible that the same value of δ would perform poorly in retaining the sequences of whistles produced by the same dolphin, without separating them into different sequences.

In general, there are several approaches in the literature that are used to label the start and end of sequences in signals produced in non-human communications, but there is little agreement on this topic [Kershenbaum et al., 2016]. Since there are no standard techniques or even a priori knowledge to evaluate the resulting sequences that we obtain, the final methodology we will choose to apply is highly dependent on the tools and data we have available. In our analyses, we propose a new approach, in which we consider a wide range of values for δ. With this approach, we aimed to study the effect of δ on the resulting sequences as well as to the analyses of linguistic laws, while at the same time avoid perceptual bias if we were to define sequences "by eye" or by choosing a fixed value for δ.

In our additional analysis of the adult and infant data sets, we did not have to make decisions regarding the separation of sequences, as this process was already performed by researchers in previous studies [McCowan et al., 1999]. For the case of infant dolphins, we do not have enough information to draw a conclusion. Even though we obtain a negative relationship between the sequence size and the whistle duration, the confidence intervals of the estimate were wide up to the range of positive values. Perhaps the absence of the law in infants is related to the fact that, like humans, they are still developing their communication. On the other hand, we do find supporting evidence for the law in adult dolphins. We consider that this evidence confirms the soundness of the results that we obtain with the large data set and supports further the adherence of dolphin whistle sequences to Menzerath's law.

Our results are in accordance with the findings of previous studies that show support of Menzerath's law in other non-human communication systems [Gustison et al., 2016, Clink et al., 2020, Heesen et al., 2019, Favaro et al., 2020]. This analysis provides the first evidence of the linguistic law in the signal communication in the family of cetaceans. These findings provide additional insights into the exploration of the properties of communication. Finally, the results can also provide valuable insights in the discussions around compression as a general principle of animal behavior [Ferrer-i-Cancho et al., 2013].

3.3 Methods

Sequences

We defined sequences using the information we have about the silent gap between the whistles. We considered 100 values in the range of 0.1 to 10 seconds with the aim to study the effect of different intervals δ. In order to construct the sequences, we followed the steps presented in Algorithm 1. The algorithm receives as input the max silent gap δ that we want to allow between two consecutive whistles of the same sequence, the list of whistle durations and the list of silences between the whistles. The core idea is that if the silence value that we are comparing in a given iteration is lower or equal to δ, then the two whistles that the silence correspond to belong to the same sequence.

Algorithm 1 buildSequences(δ, W, S)

```
Input δ: maximum silence in two consecutive vocalizations of the same sequence
Input W: list of all whistle durations of size $N$ ($N$=total number of whistles)
Input S: list of silences between the whistles of size $N-1$
sequences ← empty matrix to store sequences
current_sequence ← empty list to store whistles of current sequence
for $i \leftarrow 1$ to length(S) do
    if $S[i] \leq \delta$ then
        current_sequence ← append($W[i]$)
    else
        current_sequence ← append($W[i]$)
        sequences ← current_sequence
        current_sequence ← restore to empty list
    end
end
return sequences
```

We illustrate how this algorithm works with the means of an example. Assume a list W of size N_W=10 which contains the whistle durations, and a list S of size N_S=9 which contains the silences between the whistles. For this example, we consider a gap of δ=0.3 as the maximum silent gap between whistles of the same sequence. Finally, let's assume an empty set F which will contain all the final sequences for this δ.

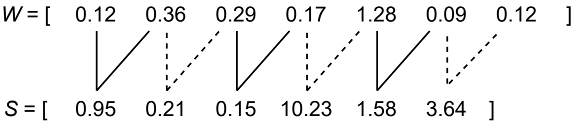


Figure 9: Example lists of whistle durations and their corresponding silences. W contains example whistle durations in order of appearance, and S the silent gaps that corresponds to each pair of whistles.

A value in S corresponds to the silence between two consecutive whistles in W. Therefore, the size of S is smaller from W by one element. In our example in figure 9, s_1=0.95 seconds is the silent gap between whistles w_1 and w_2 which have durations 0.12 and 0.36 respectively. In the first iteration, we compare if $s_1 \leq \delta$. The condition does not hold, as 0.95 >0.3, therefore w_1 is stored as a sequence of length 1, $F = [(w_1)]$. In the second iteration, the condition holds ($0.21 \leq 0.3$), so the current sequence contains w_2. For $i = 3$, again the condition holds ($0.15 \leq 0.3$), therefore we append in the sequence the current whistle, (w_2, w_3). In the case of $i = 4$, we have 10.23 >0.3, hence we stop the sequence here. We append w_4 to the current sequence, as this whistle's gap with the previous whistle is s_3 that we found to be less than δ, (w_2, w_3, w_4). We store the sequence in F. For the last two iterations, $i = 5$ and $i = 6$, the silence in both cases is greater than δ, and as a result the remaining whistles are stored in separate sequences. The list of the final sequences has the following form: $F = [(w_1), (w_2, w_3, w_4), (w_5), (w_6), (w_7)]$.

For the randomized sets of sequences, we repeat the process using the same algorithm. For each shuffling scenario, we randomize some of the algorithm's inputs. In the first case, we shuffled the order of the silences but maintained the list of whistle durations as it was. The resulted sequences here are different from those of the initial sets. In the second case, we shuffled the order of the whistles duration but kept the original order of the silences. The results here produce sequences that correspond to the initial ones in terms of size, but the whistles that they are consisted of are different. Finally, in the third case, we shuffled the order of both the silences and the durations of the whistles. For all cases, including the initial, we used the same values for δ. As a result, we produced 400 sets of sequences, to which we applied our analyses. In appendix D are included additional examples that illustrate how algorithm 9 works and how the different shufflings affect the output sequences.

The processes we describe here were performed in a Python environment, version 3.7.3. Complementary packages we used for the implementation of the algorithm and for producing summary statistics and figures are: NumPy [Harris et al., 2020], pandas [Wes McKinney, 2010] and SciPy [Virtanen et al., 2020].

Statistical Analysis

To analyze the relationship between whistle duration and sequence length (i.e. the number of whistles that belong to a sentence) we firstly applied Spearman rank correlation tests. We choose this method for being non-parametric and to remain as agnostic as possible about the exact functional dependency between the variables [Ferrer-i Cancho et al., 2014]. We consider a negative correlation between the two to be consistent with Menzerath's law. However, a negative correlation does not provide sufficient evidence that the law is present in dolphins communications. The correlation test was applied using the mean duration of a sequence, which guarantees that we have exactly one value of sequence length for one value of mean duration. In the alternate scenario, where individual durations would be used, then it would correspond multiple times the length of a sequence to the whistles that form it. Although the use of mean duration guarantees against this pseudo-replication, the aggregation of the durations can lead to susceptible results due to hidden underlying structures of the data.

To further test for Menzerath's law we constructed linear mixed models (LMMs). We ran the LMMs in R, version 3.6.3, using the function *lmer* of the R package lme4 [Bates et al., 2015]. For all sequences that corresponded to a specific δ we fitted respectively the model and compare it with its null variant. All models included the whistle duration as the outcome. The full model included as predictors the whistle's position in the sequence and the sequence length. For random effect we used the recordings id. The null model contained only the random effect without any predictors.

For all models, we tested for correlation of the fixed predictors by testing for collinearity using the variance inflation factors. For this test we used the function *vif* from the R package car [Fox and Weisberg, 2019]. The results indicated that collinearity did not exist among the predictors (VIF ≤ 3 for all cases) for all the data we used.

We followed the same methodology for the analysis of the additional data of adults and infant dolphins. The full and null models were fitted separately for the adult and infant whistle data. As the data contained information about the id and sex, we used these two as random effects both for the full and null models. The fixed effects for these cases were the same as before, the position and sequence length. Finally, for all analyses, we assessed the performance of the models against their null model using the Akaike Information Criterion (AIC). If a full model had a worse AIC value (bigger) than the null model, then we considered the model's result unreliable. For all model outputs, we have included their results along with the difference in AIC (AICd = AIC null - AIC full) in the appendix C.

4 | Revisiting clustering of dolphin whistles

Classifying types of vocalisations in animal repertoires has been a widespread problem and remains an open research field. Identifying a signal type is crucial, as it provides a mean to study the vocal signals of an individual, as well as the communication within a group and across different social populations [Anikin et al., 2018].

The majority of methods that have been developed so far base their analysis on acoustic features that are extracted from the signal's spectrogram [Kershenbaum et al., 2016]. One approach has been the visual classification of the spectrograms "by eye", which relies on the high ability of humans for pattern recognition. In this approach, the signals are subjectively characterised by a trained observer [Janik, 1999]. However, as this classification process is based on unspecified features that are meaningful to human perception, it cannot be used as a standardised method, neither it is possible to replicate its results. Additionally, the visual inspection is time-consuming and prevents its applicability in large data sets of signals.

More recent approaches have been the classification of manually extracted features of the signals (e.g. mean frequency, range, duration) using quantitative techniques [Adi et al., 2008, Kohlsdorf et al., 2014], fully automated systems for feature extraction, and classification or clustering algorithms [Kershenbaum et al., 2016]. These approaches allow the deployment of automatic analysis, taking advantage of large quantities of data in a short time. On the drawbacks, the validation of their results is a non-trivial problem, as no gold standards are available.

Particularly in the case of dolphin whistles, there are previous works that have pursued this goal. However, they either deal with very small data sets [Kershenbaum et al., 2013, McCowan, 1995] or use very simple distance metrics [McCowan, 1995]. Additionally, in the studies where the clustered whistles were compared against "ground truths", those were the signal types of a classification performed either by eye or by another clustering technique which was considered to have yielded satisfactory results. Therefore, the true whistle types that are used for assessing the quality of the clustered types are "true" according to a specific perspective. As a result, the goal of the clustering is to group together signals that have similar biological features. The results might not correspond accurately to some ground truths, but they still might have found groupings of vocalisations that are relevant to the dolphins.

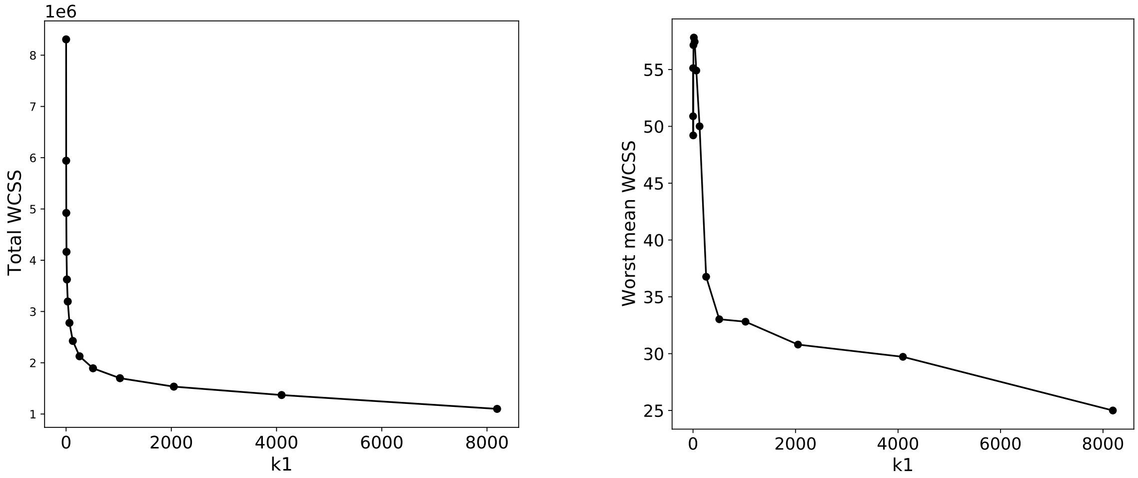
For the purposes of separating dolphin whistles into types for our work, we incorporated the methods developed by Caio Seguin for large time series data sets [Seguin, 2015]. The proposed work is a two-phase clustering methodology designed to cluster large sets of time series without the need for big computational resources. The method was tested on the same dolphin whistles data set that we described and used earlier in chapters 2 and 3.

In the present study, we replicate this work verifying its promising results, and later use them to generate the input for our quantitative linguistic research. The methodology consists of applying a two-level clustering algorithm. In the first level, the clustering algorithm is applied on the data with the goal of obtaining a compressed, representative version of the initial large data set in the form of whistle prototypes. Then, in the second phase, it computes a distance matrix of the prototypes using dynamic time warping (DTW) as a distance measure. Finally, it applies hierarchical clustering on this distance matrix using Ward's criterion as a linkage method to form the final groupings.

4.1 Results

For the first phase of the methodology, a k-Means clustering was applied to the whistle time series data, which were previously pre-processed by applying a z-score standardisation. As outlined earlier, the goal of the first phase is to compress the initial large data set into a smaller number of prototypes k_1. In order to decide a suitable value for this parameter, the phase one was repeated multiple times using a range of values for k_1 that started from 2 up to 8129 by powers of 2. To determine which are the optimal values for k_1, we analysed the results of the total WCSS (Within-Cluster Sum of Squares of the distances). Figure 10a shows that the total WCSS against the input values of k_1 forms an elbow within the interval $256 \leq k_1 \leq 1024$.

In figure 10b is depicted an additional metric to assist in the evaluation of the k_1 parameter. As in the original work, we calculated the mean WCSS for each output cluster (WCSS divided by the number of elements) and determined the worst cluster in a partition by selecting the one with the maximum value. Figure 10b shows that the curve for this metric decreases up to $k_1 = 2048$, and afterwards yields only slight improvements.



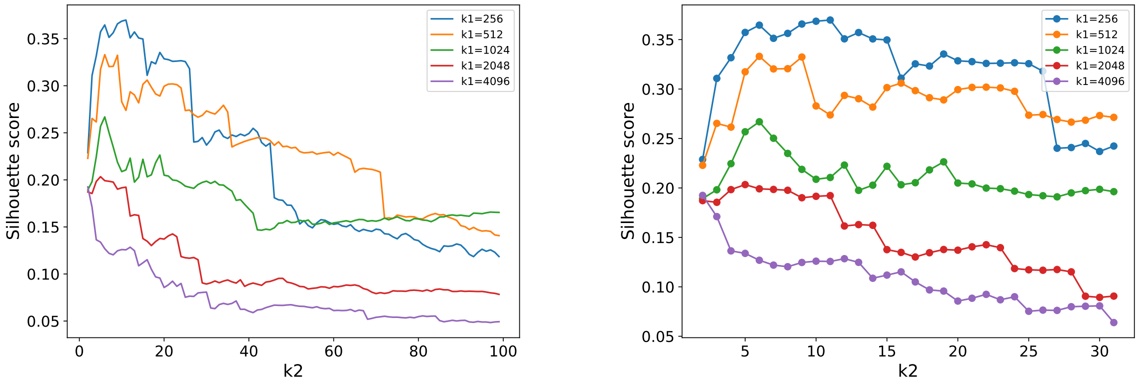
(a)

(b)

Figure 10: Total WCSS and mean worst WCSS as k_1 increases. Figure *(a)*: The total WCSS curve bends between 256 and 1024. Figure *(b)*: The mean worst WCSS curve bends close to 1024.

Moving to phase two, we calculated the DTW distance matrix between the k_1 prototypes that were found during phase one for the values in the range of $[256, 4096]$. Then, each matrix was used as an input to the final clustering method, which uses Ward's hierarchical linkage. Typically, the results of the linkage are depicted in a dendrogram which represents the nested clusters. However, due to the size of our input and the type of data, such visualizations do not provide valuable insights for our analysis. Nevertheless, refer to appendix E for resulting dendrograms.

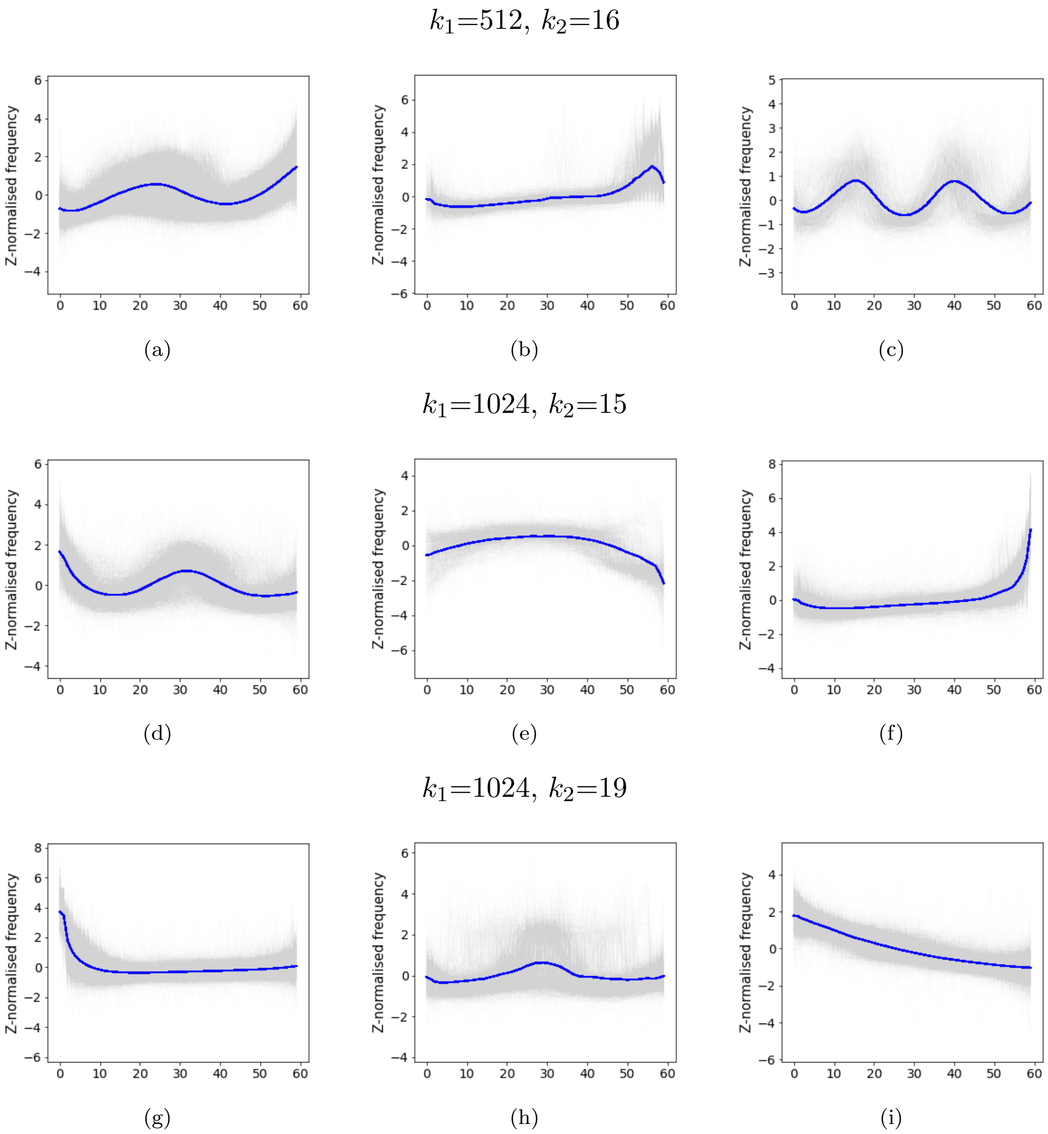
In hierarchical clustering, it is not required to pre-define the final number of output clusters. After obtaining the hierarchy of clusters as a dendrogram, the tree can be cut at various levels, depending on the needs of the analysis. In our case, to evaluate the quality of the clusters, we calculated the silhouette score for a number of possible final outputs. We cut the tree in levels that returned a final number of clusters in the range $k_2 = [2, 100]$. Figure 11a shows the silhouette scores for all k_2, and in figure 11b depicts the same results focusing in the range $[1, 30]$, where we obtained the maximum values.



(a) (b)

Figure 11: Silhouette scores for values of k_1 and k_2. In plot *(a)*, we show the results for a range of 100 values for k_2, while in *(b)* same curves as in (a) focusing on the local maxima in the values up to 30.

A complementary evaluation for the output clusters is to look into the visualizations of the grouped whistle contours. Plotting the obtained clusters can provide a broader overview of the result and determine up to a certain extent if the results make sense. Figure 12 includes plots of some example clusters of the obtained results. Each row includes three example clusters from several possible cuts of the obtained trees along with a representative of the cluster. We denote each cluster using the parameters used for obtaining it as "$k_1 = a, k_2 = b$". For example, $k_1 = 512, k_2 = 13$ are the outputs for using $k_1 = 512$ as a compression parameter and $k_2 = 13$ is the final number of clusters.



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 12: Example outputs from final clusters for combinations of k_1, k_2. The first row includes 3 example clusters (a,b and c) out of the total clusters obtained using $k_1 = 512$ and $k_2 = 16$. Similarly, the subfigures d,e and f are examples clusters of $k_1 = 1024, k_2 = 15$ and g,h and i of $k_1 = 1024, k_2 = 19$. Each a chosen representative of the cluster in blue along with the original whistle contour data that it grouped together.

4.2 Discussion

The choice of the initial number of clusters k_1 in phase one reflects the level of compression we apply to the data. As outlined in the work of Seguin [Seguin, 2015], an adequate value for k_1 needs to satisfy a number of objectives that are summarized as follows:

1. Identify groups of naturally aligned time series with the same shape, compressing only the series of similar shape into a given centroid.
2. Reduce time and memory complexity of phase 2.
3. Filter out the noise of original observations, revealing “smooth shape patterns” in the

data.

4. Improve the quality of the clustering of phase 2.

These points serve as a guide to the decision of the k_1 values that are going to be used. The obtained outputs from phase one show that after $k_1 \simeq [512, 1024]$ we get only slight improvements in the WCSS (fig. 10a). Choosing a value greater than this would have a considerable cost in the objective of time and memory reduction, as well as in the goal of filtering out noise. The results of the mean worst WCSS, on the other side, provide some additional insight into the performance of the clusters overall (fig. 10b). The essence of this metric is to indicate when the least performing cluster for a given k_1 starts to obtain good results. Using as a basis the point where that cluster starts to improve, then we can assume that the rest of the clusters are also improving. The obtained curve indicates a slightly higher value, given the point where it starts to stabilize, at $k_1 = 2048$.

In phase two, we obtain a different picture for the larger values of k_1. We see that, in general, in the cases where lower levels of compression were applied (i.e. larger numbers for k_1) the obtained clusters are returning a much lower score for the silhouette curve, which indicates a poorer fit. While for the higher levels of compression, the silhouette scores returned better results (fig. 11a). In those cases, we observe as well the presence of local maxima in the curves. For example, in the case of $k_1 = 512$, the silhouette score is showing a drop for values close to $k_2 = 10$, and then it increases again (fig. 11b). This behaviour of the curve is highlighting the values of k_2 that is worth looking further into.

For all the curves, however, we see a decreasing tendency for values greater than 30. This decrease can be the result of assigning otherwise similar contour shapes to separate clusters, because the method starts to catch small variabilities in the time series from that level and on-wards. These variabilities can be, for example, shifts at the point where the peaks of a wave shape occur in the time axis. Another potential explanation could be that the hierarchy has started being sensitive to alternations that are produced from different individuals and it is splitting them into new clusters. These sort of variations can be further evaluated by looking more closely at the different levels of a particular tree and compare the obtained clusters with the neighbour tree levels.

Summarizing these observations, this methodology provides the flexibility to decide upon a combination of the k_1 and k_2 parameters depending on the obtained results as well as on the needs of the analysis we want to perform. Most importantly though, the results can be used on a multi-level basis, where the hypotheses of analysis can be tested in multiple levels of granularity of the signal types. We believe that this approach can provide more insights into the interpretations of results instead of choosing an optimal k_1, k_2 combination. This is the approach that was followed as well for the next steps of our work in studying the law of brevity. That analysis uses multiple versions of the final repertoire obtained here, using the combinations of k_1, k_2 that exhibited local maximum values in their silhouette evaluation.

4.3 Methods

Phase 1

The first phase of the methodology begins by applying a Z-score normalization to the data. Given a data set X of time series data of length t, we z-normalize a series x using equation 4. This z-normalization was applied as a pre-processing step, with the aim to assist our analysis in the later steps for the application of the distance measure [Ding et al., 2008].

$$x_{Z-score} = \frac{x - \text{mean}(x)}{\text{sd}(x)} \tag{4}$$

Following, a K-means clustering algorithm was applied on the z-normalized data set of the time series. Specifically, the applied method uses the k-means++ variation, which runs the K-means ten times with different centroid seeds and keeps the result with the best output in terms of WCSS. This variation was introduced in order to avoid getting trapped in local minima due to poor centroid initialization [Arthur and Vassilvitskii, 2006, Seguin, 2015]. The implementation of this step was performed using the scikit-learn Python module [Buitinck et al., 2013].

Algorithm 2 Clustering phase 1

```
Require: $X = N \times t$ Z-normalized matrix of time series
Require: $k_1$ $centroids, membership\_1 \leftarrow$ K-means($X, k_1$)
return $centroids, membership\_1$
```

Phase 2

In the second phase, we initially calculated the distance matrix between the output centroids of phase 1 using as a metric the dynamic time warping distance (DTW). The distance matrix was then used as an input for the last clustering phase, which was the hierarchical linkage. The criterion used for this last step was Ward's minimization criterion. Finally, we assigned the N original observations to the chosen k_2 clusters by grouping together the time series whose corresponding centroid of phase one belonged to the same final cluster.

Algorithm 3 Clustering phase 2

```
Require $centroids = k_1 \times t$ centroid matrix from phase 1
Require: $membership\_1 = N \times 1$ membership vector from phase 1
Require: $k_2$
$D \leftarrow$ DTW matrix between $centroids$
$centroids, membership\_1 \leftarrow$ K-medoids($X, k_1$)
return $membership\_final$
```

Algorithm 4 Cluster by association

```
Require $mbs\_1 = N \times 1$ membership vector from phase 1
Require: $mbs\_2 = k_1 \times 1$ membership vector from phase 2
$mbs\_final \leftarrow membership\_1$
for $i \leftarrow 1$ to length($mbs\_2$) do
    $mbs\_final[mbs\_1 == i] \leftarrow mbs\_2[i]$
end
return $mbs\_final$
```

5 | Zipf's law of abbreviation

Zipf's law of abbreviation, also known as law of brevity, posits that the more a word is used, the shorter it tends to be [Zipf, 1949]. This qualitative statement was originally tested in texts or manuscripts, where the length of the word was measured using the number of characters the word was consisted of. More recent studies have shown that the same statistical law holds when we measure the word size using physical features, such as the duration [Tomaschek et al., 2013, Gahl, 2008, Torre et al., 2019]. Additionally, the law has been found to hold in the vast majority of languages examined [Bentz and Ferrer-i-Cancho, 2016].

As in the case of Menzerath's law, various studies are examining the existence of the law of brevity in non-human communication systems. This law is translated as a negative relationship between the frequency of use of a specific type and its mean duration. So far, this type of relationship has been found to exist in the vocalisations of Formosan macaques [Semple et al., 2010], in bats' short-range communication [Luo et al., 2013] and in penguins vocalisation [Favaro et al., 2020]. Additionally, patterns consistent with the law have been found in other forms of communication apart from signal; in the gestural signals of chimpanzees [Heesen et al., 2019] and in the repertoire of dolphin surface behaviours [Ferrer-i-Cancho and Lusseau, 2009]. For all these studies, the types of the signals were assigned either by visual evaluation of the researchers ([Semple et al., 2010]), clustered based on the spectra contour or spectrotemporal features ([Luo et al., 2013], and [Favaro et al., 2020] respectively), or recorded by an observer in the cases of gestural communication studies.

However, Zipf's law of abbreviation was not found in the studies of some species. It was not found in the vocal repertoires of common marmosets, and golden-backed uakaris [Bezerra et al., 2011] nor in the complex phrases of gibbons [Clink et al., 2020]. Additionally, there were studies that conformity to the law depended on different variables. In the study of chickadees calls, adherence to the law was dependent on the level of analysis [Hailman et al., 1985], while in a study on rock hyraxes, support was found in males but not females [Demartsev et al., 2019].

For our analysis of Zipf's law of abbreviation, we aim to investigate whether the law also holds in dolphin whistle types. Particularly for the whistle types, we use the outputs obtained from the two-phase clustering described in the previous chapter. For this analysis, we repeated our experiments on multiple combinations of the parameters k_1, k_2, each of which produces a different version of the final repertoire with k_2 final types.

5.1 Results

Initially, we looked into the relationship between the probability of occurrence of whistle type and its mean duration using correlation tests. To assess in more detail the effect of the frequency on the duration of a whistle, their relationship was explored using generalized linear mixed-effects models.

Frequency and duration of whistle types

Table 8 summarizes the results of the Spearman correlation tests for all possible whistle clustering outputs we tested for. The columns include k_2, k_2 as an identifier for the name of the whistles types output from the clustering phase, where k_1 is the parameter used for compression and k_2 is the total number of types we obtained (see chapter 4 for more details). The other two columns include ρ for the Spearman rank correlation between frequency and mean whistle type duration, and lastly, the p-value.

With the correlation tests we do not find a negative correlation in none of the clustered whistle types, with the exceptions of $k_1 = 1024, k_2 = 18$ and $k_1 = 512, k_2 16$. On the other hand, the p-values obtained for all outputs are way above any significance level, with the exception of $k_1 = 1024, k_2 = 6$ where the p-value is slightly lower than 5%.

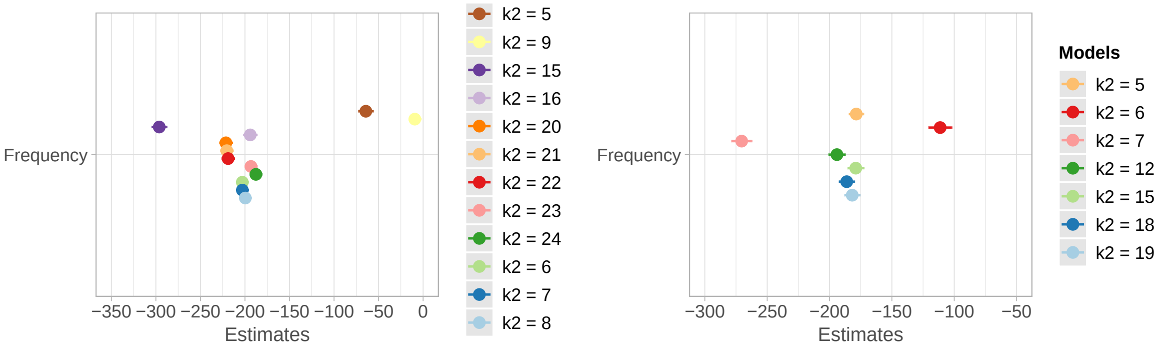
k_1, k_2	ρ	p-value	k_1, k_2	ρ	p-value
512, 5	0.700	0.1881	512, 23	0.096	0.6635
512, 6	0.372	0.4685	512, 24	0.167	0.4355
512, 7	0.072	0.8791	1024, 5	0.700	0.1881
512, 8	0.192	0.6514	1024, 6	0.829	0.0416*
512, 9	0.083	0.8312	1024, 7	0.321	0.4821
512, 15	0.047	0.8695	1024, 12	0.280	0.3786
512, 16	-0.029	0.9139	1024, 15	0.289	0.2957
512, 20	0.111	0.6405	1024, 18	-0.028	0.9126
512, 21	0.057	0.8057	1024, 19	0.095	0.6997
512, 22	0.024	0.9146			

Table 8: Spearman correlations and p-values of whistle type's frequency and mean whistle duration per type. With asterisk we highlight the values that return a p-value below the significance threshold of 0.05.

LMM statistical analysis

The linear mixed models' analyses returned consistent results regarding the effect of the frequency of use of the whistle type on the whistle duration. Figure 14 shows the estimated coefficient of frequency for all models, one for each repertoire we created. In table 10 are included detailed summaries for all models that contain the coefficient estimates, 95% confidence intervals and the AIC difference between the full and null model.

All the fitted models returned negative coefficient estimates for the frequency of whistle type. We considered all the predictors reliable, as the 95% confidence interval results did not cross zero for any of the models. Finally, all the models were ranked highest as indicated by the AIC comparison with their corresponding null models. The negative relation between the predictor and the estimated outcome indicated by these negative coefficients is a result consistent with the law of brevity.



(a) $k1 = 512$

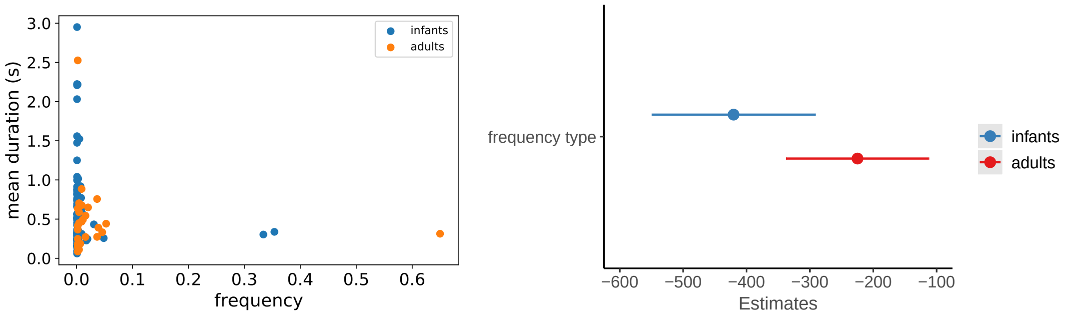
(b) $k1 = 1024$

Figure 13: Frequency estimates for all LMM models. Subfigure *(a)* include the model estimated for the repertoires we created with parameter $k1 = 512$ and subfigure *(b)* the ones created with $k1 = 1024$.

Analysis on adult and infant dolphins data

The Spearman correlation tests yielded a non-significat positive correlation for the adult dolphins (ρ=0.222, p-value=0.266), providing no evidence for patterns consistent with the law of brevity, while for infants we do obtain a negative correlation (ρ=-0.037, p-value=0.722). For both cases, however, the p-value characterises the results again as not significant.

In the case of the linear models' analysis though, we observe a different picture. For both the adult and infant dolphins we obtain negative coefficient estimates for the frequency of the whistle type, results which were supported as well by the confidence intervals of the estimates. We find that the top performing model according to the AIC was the one that contained frequency as a fixed effect and ID and sex as random intercepts. However, the difference with the model that did not contain sex as a random effect according to AIC was very small, indicating that the differences in the results due to the sex were not significant.



(a)

(b)

Figure 14: Figure *(a)*: scatter plot of mean whistle duration of type against frequency of use of type for adult and infant dolphins. Figure *(b)*: plot of linear mixed-effects model coefficients for adult and infant dolphins.

			Confidence intervals		
	Predictors	Estimate ± SE	Lower	Upper	AICd
adults	Intercept	496.57 ± 53.45	389.23	606.87	13.19
	Frequency	-224.86 ± 57.39	-339.89	-115.17	
infants	Intercept	471.13 ± 30.58	409.17	530.59	37.77
	Frequency	-420.39 ± 65.98	-550.14	-291.91	

Table 9: Results of LMM used to test for relationships between whistle duration and frequency of whistle type. AICd refers to the difference of the Akaike criterion if we subtract the AIC of the full model from the null.

5.2 Discussion

The results we obtain when we tested if the law holds by means of a correlation test between the frequency and mean duration do not show support for a negative relationship, with the exception of one case. However, the significance of these results is not supported by the returned p-value. The same results hold for the case of adult and infant dolphins analysis, where we did obtain negative correlations. The not significant p-values we obtained are probably due to the very small size of observations that we are comparing, and raise the concern of potential statistical artifacts in the results. The test was performed between the frequency of a type and the mean duration of all whistles that were clustered as this type. So, for example, in the repertoire version of $k_1 = 512, k_2 = 15$, the size of the sample in the correlation test was 15. Overall, this analysis does not provide us sufficient results to draw conclusions.

On the other hand, the linear mixed-effects model analysis uses all the observations available, as these models are able to control for repeated measurements. The results here showed a strong negative relationship between the frequency of a type and the individual duration of whistles. The relationship pattern retrieved here was the same for all the versions of repertoires used, as well as in the analysis of adult and infant dolphins data, which provides further support for adherence to the law of brevity. The evidence of the law across all the combinations of k_1, k_2, according to these models, rules out the possibility of a statistical artifact.

We obtain consistent results across all the different versions of the repertoire that we created, which arguably strengthens the hypothesis that dolphin whistles conform to this linguistic law. As we discussed previously in chapter 4, there are no known whistles classes that we can use to evaluate the clusters, and the robustness of such methodologies is a matter of debate [McCowan and Reiss, 2001, Janik, 1999]. However, having an automatic clustering technique is useful for studying such complex communication systems.

The way that the experiments were performed was affected by the lack of known whistle types. For this reason, to obtain better results that can provide more insight, we created different output repertoires based on purely machine leaning arguments. Therefore, as the output clusters are results of different cuts of the hierarchical tree that was obtained in the second phase of the clustering methodology, repertoires such as $k_1 = 1024, k_2 = 18$ and $k_1 = 1024, k_2 = 19$ will have differences in only some of their grouped types. The persistent evidence across all the repertoire versions indicate that the results are robust against small variabilities that might occur in the grouped types.

Finally, our results are consistent with the studies across diverse taxa that find patterns consistent with the law of brevity. In contrast with the study on rock hyraxes [Demartsev et al., 2019], our results did not show sex-specific conformity to the law. To conclude, these findings, along with our results on Menzerath's law, show compelling evidence of adherence to linguistic laws in the communication system of dolphins.

5.3 Methods

Whistle types

For all whistles included in the data set of 160k contours, we distinguished different types of whistles by applying the two-phase clustering technique described in chapter 4. For the detailed description of the steps followed according to it, we refer the reader to that chapter. For the k_1 and k_2 parameters of the methodology, there were used values that satisfy the clustering objectives and that according to the evaluation process performed in the previous chapter, their obtained clusters yield adequate results. Specifically, the final values used were: $k_1 \in 512, 1024$. The parameter values for k_2 were the ones that appeared to score local maxima values in figure 11b in the previous chapter. For k_2, these were 5, 6, 7, 8, 9, 15, 16, 20, 21, 22, 23, 24, while for $k_2 = 1024$, we selected the following: 5, 6, 7, 12, 15, 18, 19. In total, we used 19 possible repertoire outputs.

For the case of adult and infant whistle data sets, the individual dolphin contours had already been assigned to types. Details about the used technique can be found in [McCowan, 1995, McCowan and Reiss, 1995a, McCowan and Reiss, 1995b].

Statistical analysis

To test the hypothesis of the brevity law, correlation tests on frequency versus duration per type were initially applied. For each set of repertoires that was constructed in the clustering phase, we calculated the frequency of appearance and the mean duration of each type. The frequency of use for each call type was calculated as the number of times the type appeared out of the total number of calls. Similarly to the previous analysis in chapter 3, the Spearman rank correlation test was performed using the statistical function *spearmanr* of the scipy Python module [Virtanen et al., 2020].

The same methodology for this first step of the analysis was applied on the data of adult and infant dolphins. We note that for the calculation of mean whistle duration per type there were used fewer tokens compared to the calculation of frequency due to missing values in the field of whistle durations. In total, for the adults the missing values were 28, while for the infants 140.

We continued our analysis by building linear mixed effect models (LMM) using the *lme4* package in R [Bates et al., 2015]. The models included the individual whistle duration as the response variable and the frequency of its type as a fixed factor. As a random effect, they included the ID of the recordings. Additionally, the null model was built using the same random effect without any predictors. For evaluating the performance of the models, we used the Akaike Information Criterion. These steps were repeated for all the whistle repertoires that we clustered separately.

For the case of adult and infant data sets, four linear mixed-effects models were fitted, all of which had as an outcome the duration of the individual whistle. The first model included as fixed effect the frequency of the whistle's type, and as random effects the ID of the dolphin and its sex. The second model contained the same fixed predictors but only ID as a random effect. Finally, two null models that contained no predictors were fitted, one with the ID and sex for the random intercept and one with sex only, accounting as the null models for each of the full models described earlier. These models were fitted separately for adults and infants. For both data sets, fitting the full model that considered the two random effects, returned zero variance for the random effect of sex, leading us to the decision to fit the same model without considering it in the random effects.

			Confidence intervals		
k_1, k_2	Predictors	Estimate ± SE	Lower	Upper	AICd
512, 5	Intercept	358.681 ± 5.330	348.196	369.141	219.28
	Frequency	-64.281 ± 4.313	-72.735	-55.804	
512, 6	Intercept	342.621 ± 5.088	332.606	352.606	4.56
	Frequency	-9.296 ± 3.625	-16.400	-2.178	
512, 7	Intercept	404.015 ± 5.653	392.883	415.107	4510.34
	Frequency	-296.227 ± 4.376	-304.802	-287.636	
512, 8	Intercept	378.729 ± 5.389	368.116	389.303	2329.85
	Frequency	-194.078 ± 4.003	-201.921	-186.223	
512, 9	Intercept	381.388 ± 5.395	370.762	391.973	3402.49
	Frequency	-221.389 ± 3.773	-228.782	-213.987	
512, 15	Intercept	374.709 ± 5.393	364.084	385.290	3470.67
	Frequency	-220.392 ± 3.718	-227.678	-213.096	
512, 16	Intercept	374.187 ± 5.387	363.574	384.755	3490.67
	Frequency	-218.936 ± 3.683	-226.153	-211.709	
512, 20	Intercept	369.020 ± 5.322	358.534	379.463	2925.04
	Frequency	-193.241 ± 3.554	-200.205	-186.267	
512, 21	Intercept	367.851 ± 5.306	357.397	378.260	2813.10
	Frequency	-187.724 ± 3.521	-194.625	-180.816	
512, 22	Intercept	369.713 ± 5.322	359.227	380.155	3372.76
	Frequency	-202.965 ± 3.474	-209.773	-196.148	
512, 23	Intercept	369.683 ± 5.323	359.195	380.126	3374.11
	Frequency	-202.908 ± 3.473	-209.713	-196.095	
512, 24	Intercept	369.078 ± 5.315	358.605	379.505	3290.86
	Frequency	-199.585 ± 3.459	-206.363	-192.799	
1024, 5	Intercept	396.954 ± 5.717	385.698	408.172	3427.82
	Frequency	-178.564 ± 3.030	-184.502	-172.613	
1024, 6	Intercept	365.475 ± 5.338	354.969	375.950	536.66
	Frequency	-111.288 ± 4.787	-120.667	-101.888	
1024, 7	Intercept	396.820 ± 5.598	385.794	407.803	4079.46
	Frequency	-270.470 ± 4.204	-278.708	-262.217	
1024, 12	Intercept	374.584 ± 5.369	364.007	385.117	3116.51
	Frequency	-194.071 ± 3.457	-200.845	-187.288	
1024, 15	Intercept	370.617 ± 5.400	359.978	381.212	2788.70
	Frequency	-178.973 ± 3.371	-185.579	-172.356	
1024, 18	Intercept	370.124 ± 5.388	359.507	380.695	3323.18
	Frequency	-186.211 ± 3.211	-192.503	-179.910	
1024, 19	Intercept	369.239 ± 5.379	358.640	379.791	3198.77
	Frequency	-181.812 ± 3.196	-188.075	-175.540	

Table 10: LMM table results for each repertoire version used. Column k_1, k_2 indicates the parameters used during the clustering phase to obtain it that repertoire. The table includes as well the predictors, 95% confidence intervals and the AIC difference between the null and full models (AIC null - AIC full).

6 | General discussion and conclusions

In this work, we have studied the communication system of dolphins through the lens of quantitative linguistics. Our results show that their complex usage of vocalisations reveals statistical patterns that are consistent with Menzerath's law and Zipf's law of brevity. These findings suggest a parallel between the communication system of dolphins to that of humans, from which the linguistic laws initially were derived from, and subsequently imply conformity to the information theoretic principle of compression.

The analysis of Menzerath's law addressed the challenges of defining an appropriate threshold for segmenting vocalisations into sequences. In this study, it was introduced for the first time a methodology that studies a range of values for the threshold, and the results obtained in chapter 3 suggest that indeed the final results are heavily dependent by the choice of its value. We consider that using this method with a range of values that are meaningful for the species in study can provide solid evidence about the existence or not of patterns consistent with Menzerath's law. For the case of dolphins, it was shown that adherence to the law was found in units defined for small threshold values.

Furthermore, we introduced a new method to control for secondary properties that might lead to patterns consistent with this law. Along with the initial sets of sequences that were tested for adherence to Menzerath's law, three additional sets of shuffled sequences we created to test whether the negative relationship between construct and constituent is indeed inevitable. This way, we created control cases to compare the obtained results against the trivial hypothesis. The findings defy a trivial explanation, while the methodology we introduce here for the first time provides a new mechanism to evaluate the fit of a trivial explanation.

In the analysis performed for Zipf's law of abbreviation, we find that dolphin whistles duration has a negative relationship with the frequency of use of the whistle type. As discussed in the chapter of this law (chapter 5), a problem that remains open in the studies of animal behavior is the signal classification into types. Our approach in conducting the investigation of the law gave us insights about the statistical patterns between frequency and duration across meaningful cluster numbers that were defined during the two-level clustering method. The consistent results obtained across all of them lead us to the conclusion that the relationships obtained are robust against changes in the final cluster number that was defined. However, we do not know to what extend the clustering application affects the results overall. Therefore, it is worth to question, if the methods for classifying the whistles into types was improved, if we would we still find the negative relationship dictated by the law in these results.

An additional remark related to the whistle types has to do with the general context in which the recordings of the large data set we worked with were collected. As described in chapter 2, the data collection was performed by recording for a certain period four dolphins that were kept in captivity. This has allowed the researchers to build a very large data set of whistles, however we do not know to what extent the environment of the dolphins affect their vocalisations. An interesting future direction would be to examine the complex signals of dolphins in the wild, whose repertoire might include the representation of more concepts as a result of the diverse stimuli that their natural environment provides.

The existence of the two linguistics laws in the communication system of dolphins provides compelling evidence and insights in the discussion of linguistics universals, and in particular for compression as a general principle of animal behaviour. In spite of the strikingly different natural habitat of dolphins, we find the same abstract cost minimization principles are present across different species and taxa, including humans. The present findings should motivate further investigation in the complex forms of communication that we find in our close - and not so close - evolutionary relatives, as they can help in understanding the forces that drive and constrain the evolution of communication, and ultimately the evolution of languages.

References

[Adi et al., 2008] Adi, K., Sonstrom, K. E., Scheifele, P. M., and Johnson, M. T. (2008). Unsupervised validity measures for vocalization clustering. In *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 4377–4380. IEEE.

[Altmann, 1980] Altmann, G. (1980). Prolegomena to Menzerath's law. *Glottometrika*, 2(2):1–10.

[Anikin et al., 2018] Anikin, A., Bååth, R., and Persson, T. (2018). Human non-linguistic vocal repertoire: Call types and their meaning. *Journal of nonverbal behavior*, 42(1):53–80.

[Arthur and Vassilvitskii, 2006] Arthur, D. and Vassilvitskii, S. (2006). k-means++: The advantages of careful seeding. Technical report, Stanford.

[Au, 1993] Au, W. W. (1993). *The sonar of dolphins.* Springer Science & Business Media.

[Bates et al., 2015] Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

[Bentz and Ferrer-i-Cancho, 2016] Bentz, C. and Ferrer-i-Cancho, R. (2016). Zipf's law of abbreviation as a language universal. In *Proceedings of the Leiden workshop on capturing phylogenetic algorithms for linguistics*, pages 1–4. University of Tübingen.

[Bezerra et al., 2011] Bezerra, B. M., Souto, A. S., Radford, A. N., and Jones, G. (2011). Brevity is not always a virtue in primate communication. *Biology letters*, 7(1):23–25.

[Boroda and Altmann, 1991] Boroda, M. and Altmann, G. (1991). Menzerath's law in musical texts. *Musikometrica*, 3:1–13.

[Buckstaff, 2004] Buckstaff, K. C. (2004). Effects of watercraft noise on the acoustic behavior of bottlenose dolphins, Tursiops truncatus, in sarasota bay, florida. *Marine mammal science*, 20(4):709–725.

[Buitinck et al., 2013] Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., Niculae, V., Prettenhofer, P., Gramfort, A., Grobler, J., Layton, R., VanderPlas, J., Joly, A., Holt, B., and Varoquaux, G. (2013). API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pages 108–122.

[Caldwell and Caldwell, 1968] Caldwell, M. C. and Caldwell, D. K. (1968). Vocalization of naive captive dolphins in small groups. *Science*, 159(3819):1121–1123.

[Clink et al., 2020] Clink, D. J., Ahmad, A. H., and Klinck, H. (2020). Brevity is not a universal in animal communication: evidence for compression depends on the unit of analysis in small ape vocalizations. *Royal Society open science*, 7(4):200151.

[Corral and Serra, 2020] Corral, Á. and Serra, I. (2020). The brevity law as a scaling law, and a possible origin of Zipf's law for word frequencies. *Entropy*, 22(2):224.

[Cover and Thomas, 2006] Cover, T. M. and Thomas, J. A. (2006). Elements of information theory second edition solutions to problems. *Internet Access.*

[Degroot, 1986] Degroot, M. H. (1986). The negative binomial distribution. In *Probability and Statistics*, page 258–263. Addison-Wesley, Reading, MA.

[Demartsev et al., 2019] Demartsev, V., Gordon, N., Barocas, A., Bar-Ziv, E., Ilany, T., Goll, Y., Ilany, A., and Geffen, E. (2019). The "Law of Brevity" in animal communication: Sex-specific signaling optimization is determined by call amplitude rather than duration. *Evolution letters*, 3(6):623–634.

[Ding et al., 2008] Ding, H., Trajcevski, G., Scheuermann, P., Wang, X., and Keogh, E. (2008). Querying and mining of time series data: experimental comparison of representations and distance measures. *Proceedings of the VLDB Endowment*, 1(2):1542–1552.

[Favaro et al., 2020] Favaro, L., Gamba, M., Cresta, E., Fumagalli, E., Bandoli, F., Pilenga, C., Isaja, V., Mathevon, N., and Reby, D. (2020). Do penguins' vocal sequences conform to linguistic laws? *Biology letters*, 16(2):20190589.

[Fedurek et al., 2017] Fedurek, P., Zuberbühler, K., and Semple, S. (2017). Trade-offs in the production of animal vocal sequences: insights from the structure of wild chimpanzee pant hoots. *Frontiers in zoology*, 14(1):1–10.

[Ferrer-i Cancho, 2016] Ferrer-i Cancho, R. (2016). Compression and the origins of Zipf's law for word frequencies. *Complexity*, 21(S2):409–411.

[Ferrer-i Cancho et al., 2020] Ferrer-i Cancho, R., Bentz, C., and Seguin, C. (2020). Optimal coding and the origins of zipfian laws. *Journal of Quantitative Linguistics*, pages 1–30.

[Ferrer-i-Cancho and Forns, 2010] Ferrer-i-Cancho, R. and Forns, N. (2010). The self-organization of genomes. *Complexity*, 15(5):34–36.

[Ferrer-i Cancho et al., 2014] Ferrer-i Cancho, R., Hernández-Fernández, A., Baixeries, J., Debowski, Ł., and Mačutek, J. (2014). When is Menzerath-Altmann law mathematically trivial? a new approach. *Statistical applications in genetics and molecular biology*, 13(6):633–644.

[Ferrer-i-Cancho et al., 2013] Ferrer-i-Cancho, R., Hernández-Fernández, A., Lusseau, D., Agoramoorthy, G., Hsu, M. J., and Semple, S. (2013). Compression as a universal principle of animal behavior. *Cognitive Science*, 37(8):1565–1578.

[Ferrer-i-Cancho and Lusseau, 2009] Ferrer-i-Cancho, R. and Lusseau, D. (2009). Efficient coding in dolphin surface behavioral patterns. *Complexity*, 14(5):23–25.

[Fox and Weisberg, 2019] Fox, J. and Weisberg, S. (2019). *An R Companion to Applied Regression*. Sage, Thousand Oaks CA, third edition.

[Freitag and Tyack, 1993] Freitag, L. E. and Tyack, P. L. (1993). Passive acoustic localization of the atlantic bottlenose dolphin using whistles and echolocation clicks. *The Journal of the Acoustical Society of America*, 93(4):2197–2205.

[Gahl, 2008] Gahl, S. (2008). Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language*, 84(3):474–496.

[Gustison et al., 2016] Gustison, M. L., Semple, S., Ferrer-i-Cancho, R., and Bergman, T. J. (2016). Gelada vocal sequences follow menzerath's linguistic law. *Proceedings of the National Academy of Sciences*, 113(19):E2750–E2758.

[Hailman et al., 1985] Hailman, J. P., Ficken, M. S., and Ficken, R. W. (1985). The 'chick-a-dee' calls of Parus atricapillus: a recombinant system of animal communication compared with written English. *Semiotica*, 56(3-4):191–224.

[Harris et al., 2020] Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del R'ıo, J. F., Wiebe, M., Peterson, P., G'erard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585(7825):357–362.

[Heesen et al., 2019] Heesen, R., Hobaiter, C., Ferrer-i-Cancho, R., and Semple, S. (2019). Linguistic laws in chimpanzee gestural communication. *Proceedings of the Royal Society B*, 286(1896):20182900.

[Hernández-Fernández et al., 2011] Hernández-Fernández, A., Baixeries, J., Forns, N., and Ferrer-i Cancho, R. (2011). Size of the whole versus number of parts in genomes. *Entropy*, 13(8):1465–1480.

[Janik, 1999] Janik, V. M. (1999). Pitfalls in the categorization of behaviour: a comparison of dolphin whistle classification methods. *Animal Behaviour*, 57(1):133–143.

[Kershenbaum et al., 2016] Kershenbaum, A., Blumstein, D. T., Roch, M. A., Akçay, Ç., Backus, G., Bee, M. A., Bohn, K., Cao, Y., Carter, G., Cäsar, C., et al. (2016). Acoustic sequences in non-human animals: a tutorial review and prospectus. *Biological Reviews*, 91(1):13–52.

[Kershenbaum et al., 2013] Kershenbaum, A., Sayigh, L. S., and Janik, V. M. (2013). The encoding of individual identity in dolphin signature whistles: How much information is needed? *PloS one*, 8(10):e77671.

[Kohlsdorf et al., 2014] Kohlsdorf, D., Mason, C., Herzing, D., and Starner, T. (2014). Probabilistic extraction and discovery of fundamental units in dolphin whistles. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8242–8246. IEEE.

[Krützen et al., 2005] Krützen, M., Mann, J., Heithaus, M. R., Connor, R. C., Bejder, L., and Sherwin, W. B. (2005). Cultural transmission of tool use in bottlenose dolphins. *Proceedings of the National Academy of Sciences*, 102(25):8939–8943.

[Luo et al., 2013] Luo, B., Jiang, T., Liu, Y., Wang, J., Lin, A., Wei, X., and Feng, J. (2013). Brevity is prevalent in bat short-range communication. *Journal of Comparative Physiology A*, 199(4):325–333.

[Marino et al., 2007] Marino, L., Connor, R. C., Fordyce, R. E., Herman, L. M., Hof, P. R., Lefebvre, L., Lusseau, D., McCowan, B., Nimchinsky, E. A., Pack, A. A., et al. (2007). Cetaceans have complex brains for complex cognition. *PLoS Biol*, 5(5):e139.

[Markov and Ostrovskaya, 1990] Markov, V. I. and Ostrovskaya, V. M. (1990). Organization of communication system in Tursiops truncatus Montagu. In *Sensory abilities of cetaceans*, pages 599–622. Springer.

[McCowan, 1995] McCowan, B. (1995). A new quantitative technique for categorizing whistles using simulated signals and whistles from captive bottlenose dolphins (delphinidae, tursiops truncatus). *Ethology*, 100(3):177–193.

[McCowan et al., 2005] McCowan, B., Doyle, L. R., Jenkins, J. M., and Hanser, S. F. (2005). The appropriate use of zipf's law in animal communication studies. *Animal Behaviour*, 69(1):F1–F7.

[McCowan et al., 1999] McCowan, B., Hanser, S. F., and Doyle, L. R. (1999). Quantitative tools for comparing animal communication systems: information theory applied to bottlenose dolphin whistle repertoires. *Animal behaviour*, 57(2):409–419.

[McCowan and Reiss, 1995a] McCowan, B. and Reiss, D. (1995a). Quantitative comparison of whistle repertoires from captive adult bottlenose dolphins (delphinidae, tursiops truncatus): A re-evaluation of the signature whistle hypothesis. *Ethology*, 100(3):194–209.

[McCowan and Reiss, 1995b] McCowan, B. and Reiss, D. (1995b). Whistle contour development in captive-born infant bottlenose dolphins (tursiops truncatus): Role of learning. *Journal of Comparative Psychology*, 109(3):242.

[McCowan and Reiss, 2001] McCowan, B. and Reiss, D. (2001). The fallacy of 'signature whistles' in bottlenose dolphins: a comparative perspective of 'signature information' in animal vocalizations. *Animal behaviour*, 62(6):1151–1162.

[Menzerath, 1954] Menzerath, P. (1954). *Die architektonik des deutschen wortschatzes*, volume 3. F. Dümmler.

[Petrella et al., 2012] Petrella, V., Martinez, E., Anderson, M. G., and Stockin, K. A. (2012). Whistle characteristics of common dolphins (Delphinus sp.) in the hauraki gulf, new zealand. *Marine Mammal Science*, 28(3):479–496.

[Reiss and Marino, 2001] Reiss, D. and Marino, L. (2001). Mirror self-recognition in the bottlenose dolphin: A case of cognitive convergence. *Proceedings of the National Academy of Sciences*, 98(10):5937–5942.

[Reiss and McCowan, 1993] Reiss, D. and McCowan, B. (1993). Spontaneous vocal mimicry and production by bottlenose dolphins (Tursiops truncatus): evidence for vocal learning. *Journal of Comparative Psychology*, 107(3):301.

[Seguin, 2015] Seguin, C. (2015). Categorizing dolphin whistles.

[Semple et al., 2010] Semple, S., Hsu, M. J., and Agoramoorthy, G. (2010). Efficiency of coding in macaque vocal communication. *Biology letters*, 6(4):469–471.

[Shahzad et al., 2015] Shahzad, K., Mittenthal, J. E., and Caetano-Anollés, G. (2015). The organization of domains in proteins obeys Menzerath-Altmann's law of language. *BMC systems biology*, 9(1):1–13.

[Solé, 2010] Solé, R. V. (2010). Genome size, self-organization and DNA's dark matter.

[Suzuki et al., 2020] Suzuki, T. N., Wheatcroft, D., and Griesser, M. (2020). The syntax–semantics interface in animal vocal communication. *Philosophical Transactions of the Royal Society B*, 375(1789):20180405.

[Tomaschek et al., 2013] Tomaschek, F., Wieling, M., Arnold, D., and Baayen, R. H. (2013). Word frequency, vowel length and vowel quality in speech production: An EMA study of the importance of experience.

[Torre et al., 2019] Torre, I. G., Luque, B., Lacasa, L., Kello, C. T., and Hernández-Fernández, A. (2019). On the physical origin of linguistic laws and lognormality in speech. *Royal Society open science*, 6(8):191023.

[Torre et al., 2017] Torre, I. G., Luque, B., Lacasa, L., Luque, J., and Hernández-Fernández, A. (2017). Emergence of linguistic laws in human voice. *Scientific reports*, 7(1):1–10.

[Virtanen et al., 2020] Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272.

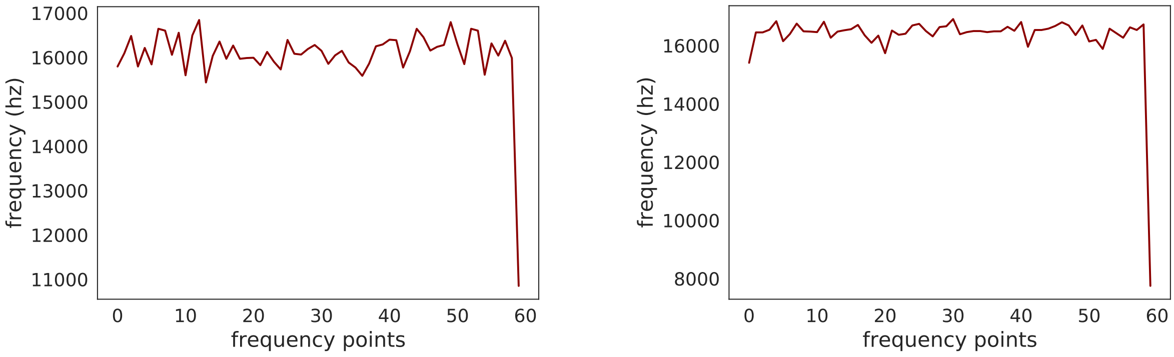
[Wes McKinney, 2010] Wes McKinney (2010). Data Structures for Statistical Computing in Python. In Stéfan van der Walt and Jarrod Millman, editors, *Proceedings of the 9th Python in Science Conference*, pages 56 – 61.

[Zipf, 1949] Zipf, G. K. (1949). Human behaviour and the principle of least-effort. Cambridge MA edn. *Reading: Addison-Wesley.*

Appendices

A | Dolphin whistles data set

Following, figure 15, on the right, includes two examples of the entries that we discarded from considering as whistle contour data. Below, in figure 16 we see the distribution of the whistle duration in the violin plots. The shape of the violin changes effectively due to the effect of the extreme values. Discarding these problematic values gives us a clearer picture of the data distribution. Note that the values of the left violin plot go up to 56 seconds, which was the maximum duration observed in the raw data. We cut the limit of the y axis for visual reasons.



(a)

(b)

Figure 15: Examples of data included in the initial data set, that do not resemble whistle contours.

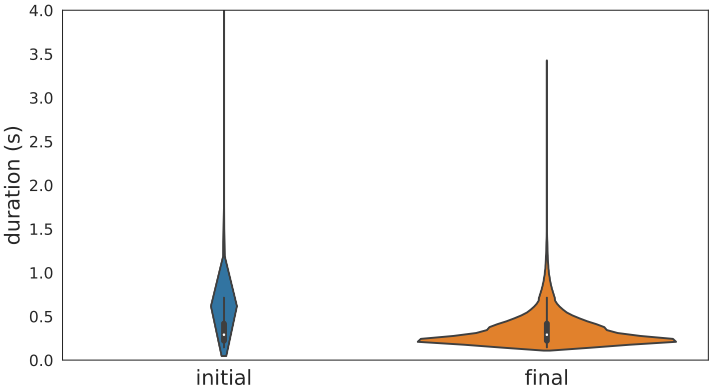


Figure 16: Violin plots of the distribution of whistle duration. *Left*: the initial data before removing the extreme entries. Note that the scale of the y-axis was cropped for visual reasons. *Right*: distribution of durations after removing problematic data.

B Sequences statistics tables

Table 11: Sequences length summary statistics - initial

δ	*Min*	*Max*	*Mean*	*St.Dev*	*N*
0.1	1	33	1.206	0.646	135328
0.2	1	192	1.298	1.251	125744
0.3	1	318	1.365	1.838	119605
0.4	1	381	1.430	2.349	114119
0.5	1	381	1.491	2.804	109431
0.6	1	518	1.551	3.493	105244
0.7	1	520	1.604	3.904	101736
0.8	1	520	1.653	4.197	98733
0.9	1	764	1.702	4.646	95921
1.0	1	1164	1.746	5.456	93503
1.1	1	1164	1.785	5.837	91418
1.2	1	1404	1.824	6.734	89494
1.3	1	1404	1.862	6.919	87646
1.4	1	1404	1.897	7.025	86034
1.5	1	1404	1.932	7.244	84470
1.6	1	1404	1.965	7.336	83065
1.7	1	1404	2.000	7.455	81630
1.8	1	1404	2.031	7.644	80370
1.9	1	1404	2.062	7.739	79167
2.0	1	1404	2.092	7.816	78016
2.1	1	1404	2.124	7.912	76867
2.2	1	1404	2.154	8.018	75773
2.3	1	1404	2.182	8.111	74808
2.4	1	1404	2.211	8.207	73841
2.5	1	1404	2.238	8.288	72933
2.6	1	1404	2.267	8.368	72006
2.7	1	1404	2.293	8.744	71175
2.8	1	1404	2.319	8.873	70383
2.9	1	1404	2.346	8.967	69576
3.0	1	1404	2.373	9.048	68775
3.1	1	1404	2.400	9.228	68002
3.2	1	1404	2.427	9.300	67264
3.3	1	1404	2.452	9.399	66583
3.4	1	1404	2.476	9.469	65918
3.5	1	1404	2.502	9.537	65232
3.6	1	1404	2.529	9.704	64553
3.7	1	1404	2.554	9.816	63906
3.8	1	1404	2.579	9.890	63294
3.9	1	1404	2.605	9.991	62659
4.0	1	1404	2.630	10.059	62072
4.1	1	1404	2.654	10.138	61512
4.2	1	1404	2.679	10.215	60938
4.3	1	1404	2.703	10.311	60389
4.4	1	1404	2.728	10.808	59848
4.5	1	1404	2.751	10.907	59335
4.6	1	1404	2.775	11.004	58829
4.7	1	1404	2.798	11.063	58338
4.8	1	1404	2.824	11.137	57814
4.9	1	1404	2.848	11.231	57317
5.0	1	1404	2.871	11.294	56852
5.1	1	1404	2.894	11.380	56411
5.2	1	1404	2.918	11.450	55940
5.3	1	1404	2.942	11.575	55490
5.4	1	1404	2.967	11.684	55016
5.5	1	1404	2.991	11.788	54574
5.6	1	1404	3.013	11.854	54170
5.7	1	1404	3.037	11.915	53756
5.8	1	1404	3.060	11.969	53353
5.9	1	1404	3.085	12.036	52911
6.0	1	1404	3.110	12.098	52485
6.1	1	1404	3.133	12.157	52107
6.2	1	1404	3.156	12.229	51725
6.3	1	1404	3.179	12.309	51346
6.4	1	1404	3.203	12.373	50970
6.5	1	1404	3.227	12.434	50589
6.6	1	1404	3.249	12.525	50240
6.7	1	1404	3.270	12.575	49922
6.8	1	1404	3.294	12.652	49564
6.9	1	1506	3.315	12.930	49244
7.0	1	1506	3.340	12.999	48875
7.1	1	1506	3.363	13.057	48545
7.2	1	1506	3.386	13.126	48206
7.3	1	1506	3.407	13.177	47908
7.4	1	1506	3.431	13.271	47580
7.5	1	1506	3.457	13.346	47218
7.6	1	1506	3.480	13.397	46909
7.7	1	1506	3.505	13.458	46578
7.8	1	1506	3.527	13.513	46289
7.9	1	1506	3.548	13.639	46006
8.0	1	1506	3.569	13.687	45735
8.1	1	1506	3.592	13.750	45445
8.2	1	1506	3.615	13.808	45159
8.3	1	1506	3.639	13.869	44862
8.4	1	1506	3.664	13.931	44558
8.5	1	1506	3.686	13.989	44291
8.6	1	1506	3.706	14.034	44051
8.7	1	1506	3.727	14.088	43797
8.8	1	1506	3.754	14.153	43489
8.9	1	1506	3.777	14.205	43218
9.0	1	1506	3.801	14.265	42948
9.1	1	1506	3.826	14.324	42672
9.2	1	1506	3.848	14.384	42425
9.3	1	1506	3.871	14.512	42172
9.4	1	1506	3.894	14.562	41921
9.5	1	1506	3.916	14.618	41688
9.6	1	1506	3.939	14.672	41439
9.7	1	1506	3.964	14.791	41183
9.8	1	1506	3.986	14.850	40955
9.9	1	1506	4.006	14.900	40746
10.0	1	1506	4.029	14.955	40516

Table 12: Sequences length summary statistics - shuffled silences

δ	*Min*	*Max*	*Mean*	*St.Dev*	*N*
0.1	1	7	1.206	0.496	135328
0.2	1	8	1.298	0.621	125744
0.3	1	12	1.365	0.707	119605
0.4	1	12	1.430	0.786	114119
0.5	1	12	1.491	0.857	109431
0.6	1	13	1.551	0.926	105244
0.7	1	13	1.604	0.984	101736
0.8	1	13	1.653	1.040	98733
0.9	1	13	1.702	1.094	95921
1.0	1	13	1.746	1.144	93503
1.1	1	15	1.785	1.188	91418
1.2	1	17	1.824	1.229	89494
1.3	1	17	1.862	1.270	87646
1.4	1	19	1.897	1.310	86034
1.5	1	19	1.932	1.345	84470
1.6	1	19	1.965	1.380	83065
1.7	1	19	2.000	1.416	81630
1.8	1	19	2.031	1.449	80370
1.9	1	19	2.062	1.482	79167
2.0	1	19	2.092	1.515	78016
2.1	1	19	2.124	1.549	76867
2.2	1	19	2.154	1.583	75773
2.3	1	19	2.182	1.612	74808
2.4	1	19	2.211	1.645	73841
2.5	1	19	2.238	1.674	72933
2.6	1	19	2.267	1.706	72006
2.7	1	19	2.293	1.737	71175
2.8	1	19	2.319	1.763	70383
2.9	1	19	2.346	1.791	69576
3.0	1	19	2.373	1.819	68775
3.1	1	19	2.400	1.849	68002
3.2	1	19	2.427	1.876	67264
3.3	1	24	2.452	1.901	66583
3.4	1	24	2.476	1.926	65918
3.5	1	24	2.502	1.950	65232
3.6	1	27	2.529	1.977	64553
3.7	1	27	2.554	2.006	63906
3.8	1	27	2.579	2.033	63294
3.9	1	27	2.605	2.060	62659
4.0	1	27	2.630	2.087	62072
4.1	1	27	2.654	2.113	61512
4.2	1	27	2.679	2.139	60938
4.3	1	27	2.703	2.163	60389
4.4	1	27	2.728	2.189	59848
4.5	1	27	2.751	2.213	59335
4.6	1	27	2.775	2.238	58829
4.7	1	27	2.798	2.261	58338
4.8	1	30	2.824	2.289	57814
4.9	1	30	2.848	2.312	57317
5.0	1	30	2.871	2.334	56852
5.1	1	30	2.894	2.359	56411
5.2	1	30	2.918	2.384	55940
5.3	1	30	2.942	2.413	55490
5.4	1	30	2.967	2.438	55016
5.5	1	30	2.991	2.460	54574
5.6	1	30	3.013	2.483	54170
5.7	1	30	3.037	2.506	53756
5.8	1	30	3.060	2.532	53353
5.9	1	30	3.085	2.557	52911
6.0	1	30	3.110	2.579	52485
6.1	1	30	3.133	2.603	52107
6.2	1	30	3.156	2.628	51725
6.3	1	30	3.179	2.651	51346
6.4	1	30	3.203	2.676	50970
6.5	1	30	3.227	2.701	50589
6.6	1	30	3.249	2.726	50240
6.7	1	30	3.270	2.748	49922
6.8	1	30	3.294	2.772	49564
6.9	1	30	3.315	2.794	49244
7.0	1	30	3.340	2.816	48875
7.1	1	30	3.363	2.838	48545
7.2	1	30	3.386	2.858	48206
7.3	1	30	3.407	2.878	47908
7.4	1	30	3.431	2.903	47580
7.5	1	30	3.457	2.932	47218
7.6	1	30	3.480	2.952	46909
7.7	1	30	3.505	2.976	46578
7.8	1	30	3.527	2.996	46289
7.9	1	30	3.548	3.019	46006
8.0	1	30	3.569	3.040	45735
8.1	1	30	3.592	3.066	45445
8.2	1	30	3.615	3.087	45159
8.3	1	30	3.639	3.110	44862
8.4	1	30	3.664	3.132	44558
8.5	1	30	3.686	3.152	44291
8.6	1	30	3.706	3.172	44051
8.7	1	30	3.727	3.190	43797
8.8	1	30	3.754	3.220	43489
8.9	1	30	3.777	3.247	43218
9.0	1	46	3.801	3.275	42948
9.1	1	46	3.826	3.302	42672
9.2	1	46	3.848	3.319	42425
9.3	1	46	3.871	3.342	42172
9.4	1	46	3.894	3.368	41921
9.5	1	46	3.916	3.392	41688
9.6	1	46	3.939	3.420	41439
9.7	1	46	3.964	3.444	41183
9.8	1	46	3.986	3.466	40955
9.9	1	46	4.006	3.484	40746
10.0	1	46	4.029	3.509	40516

Table 13: Sequences length summary statistics - shuffled durations

δ	*Min*	*Max*	*Mean*	*St.Dev*	*N*
0.1	1	33	1.206	0.646	135328
0.2	1	192	1.298	1.251	125744
0.3	1	318	1.365	1.838	119605
0.4	1	381	1.430	2.349	114119
0.5	1	381	1.491	2.804	109431
0.6	1	518	1.551	3.493	105244
0.7	1	520	1.604	3.904	101736
0.8	1	520	1.653	4.197	98733
0.9	1	764	1.702	4.646	95921
1.0	1	1164	1.746	5.456	93503
1.1	1	1164	1.785	5.837	91418
1.2	1	1404	1.824	6.734	89494
1.3	1	1404	1.862	6.919	87646
1.4	1	1404	1.897	7.025	86034
1.5	1	1404	1.932	7.244	84470
1.6	1	1404	1.965	7.336	83065
1.7	1	1404	2.000	7.455	81630
1.8	1	1404	2.031	7.644	80370
1.9	1	1404	2.062	7.739	79167
2.0	1	1404	2.092	7.816	78016
2.1	1	1404	2.124	7.912	76867
2.2	1	1404	2.154	8.018	75773
2.3	1	1404	2.182	8.111	74808
2.4	1	1404	2.211	8.207	73841
2.5	1	1404	2.238	8.288	72933
2.6	1	1404	2.267	8.368	72006
2.7	1	1404	2.293	8.744	71175
2.8	1	1404	2.319	8.873	70383
2.9	1	1404	2.346	8.967	69576
3.0	1	1404	2.373	9.048	68775
3.1	1	1404	2.400	9.228	68002
3.2	1	1404	2.427	9.300	67264
3.3	1	1404	2.452	9.399	66583
3.4	1	1404	2.476	9.469	65918
3.5	1	1404	2.502	9.537	65232
3.6	1	1404	2.529	9.704	64553
3.7	1	1404	2.554	9.816	63906
3.8	1	1404	2.579	9.890	63294
3.9	1	1404	2.605	9.991	62659
4.0	1	1404	2.630	10.059	62072
4.1	1	1404	2.654	10.138	61512
4.2	1	1404	2.679	10.215	60938
4.3	1	1404	2.703	10.311	60389
4.4	1	1404	2.728	10.808	59848
4.5	1	1404	2.751	10.907	59335
4.6	1	1404	2.775	11.004	58829
4.7	1	1404	2.798	11.063	58338
4.8	1	1404	2.824	11.137	57814
4.9	1	1404	2.848	11.231	57317
5.0	1	1404	2.871	11.294	56852
5.1	1	1404	2.894	11.380	56411
5.2	1	1404	2.918	11.450	55940
5.3	1	1404	2.942	11.575	55490
5.4	1	1404	2.967	11.684	55016
5.5	1	1404	2.991	11.788	54574
5.6	1	1404	3.013	11.854	54170
5.7	1	1404	3.037	11.915	53756
5.8	1	1404	3.060	11.969	53353
5.9	1	1404	3.085	12.036	52911
6.0	1	1404	3.110	12.098	52485
6.1	1	1404	3.133	12.157	52107
6.2	1	1404	3.156	12.229	51725
6.3	1	1404	3.179	12.309	51346
6.4	1	1404	3.203	12.373	50970
6.5	1	1404	3.227	12.434	50589
6.6	1	1404	3.249	12.525	50240
6.7	1	1404	3.270	12.575	49922
6.8	1	1404	3.294	12.652	49564
6.9	1	1506	3.315	12.930	49244
7.0	1	1506	3.340	12.999	48875
7.1	1	1506	3.363	13.057	48545
7.2	1	1506	3.386	13.126	48206
7.3	1	1506	3.407	13.177	47908
7.4	1	1506	3.431	13.271	47580
7.5	1	1506	3.457	13.346	47218
7.6	1	1506	3.480	13.397	46909
7.7	1	1506	3.505	13.458	46578
7.8	1	1506	3.527	13.513	46289
7.9	1	1506	3.548	13.639	46006
8.0	1	1506	3.569	13.687	45735
8.1	1	1506	3.592	13.750	45445
8.2	1	1506	3.615	13.808	45159
8.3	1	1506	3.639	13.869	44862
8.4	1	1506	3.664	13.931	44558
8.5	1	1506	3.686	13.989	44291
8.6	1	1506	3.706	14.034	44051
8.7	1	1506	3.727	14.088	43797
8.8	1	1506	3.754	14.153	43489
8.9	1	1506	3.777	14.205	43218
9.0	1	1506	3.801	14.265	42948
9.1	1	1506	3.826	14.324	42672
9.2	1	1506	3.848	14.384	42425
9.3	1	1506	3.871	14.512	42172
9.4	1	1506	3.894	14.562	41921
9.5	1	1506	3.916	14.618	41688
9.6	1	1506	3.939	14.672	41439
9.7	1	1506	3.964	14.791	41183
9.8	1	1506	3.986	14.850	40955
9.9	1	1506	4.006	14.900	40746
10.0	1	1506	4.029	14.955	40516

Table 14: Sequences length summary statistics - shuffled silences and durations

δ	*Min*	*Max*	*Mean*	*St.Dev*	*N*
0.1	1.0	7	1.206	0.496	135328
0.2	1.0	8	1.298	0.621	125744
0.3	1.0	12	1.365	0.707	119605
0.4	1.0	12	1.430	0.786	114119
0.5	1.0	12	1.491	0.857	109431
0.6	1.0	13	1.551	0.926	105244
0.7	1.0	13	1.604	0.984	101736
0.8	1.0	13	1.653	1.040	98733
0.9	1.0	13	1.702	1.094	95921
1.0	1.0	13	1.746	1.144	93503
1.1	1.0	15	1.785	1.188	91418
1.2	1.0	17	1.824	1.229	89494
1.3	1.0	17	1.862	1.270	87646
1.4	1.0	19	1.897	1.310	86034
1.5	1.0	19	1.932	1.345	84470
1.6	1.0	19	1.965	1.380	83065
1.7	1.0	19	2.000	1.416	81630
1.8	1.0	19	2.031	1.449	80370
1.9	1.0	19	2.062	1.482	79167
2.0	1.0	19	2.092	1.515	78016
2.1	1.0	19	2.124	1.549	76867
2.2	1.0	19	2.154	1.583	75773
2.3	1.0	19	2.182	1.612	74808
2.4	1.0	19	2.211	1.645	73841
2.5	1.0	19	2.238	1.674	72933
2.6	1.0	19	2.267	1.706	72006
2.7	1.0	19	2.293	1.737	71175
2.8	1.0	19	2.319	1.763	70383
2.9	1.0	19	2.346	1.791	69576
3.0	1.0	19	2.373	1.819	68775
3.1	1.0	19	2.400	1.849	68002
3.2	1.0	19	2.427	1.876	67264
3.3	1.0	24	2.452	1.901	66583
3.4	1.0	24	2.476	1.926	65918
3.5	1.0	24	2.502	1.950	65232
3.6	1.0	27	2.529	1.977	64553
3.7	1.0	27	2.554	2.006	63906
3.8	1.0	27	2.579	2.033	63294
3.9	1.0	27	2.605	2.060	62659
4.0	1.0	27	2.630	2.087	62072
4.1	1.0	27	2.654	2.113	61512
4.2	1.0	27	2.679	2.139	60938
4.3	1.0	27	2.703	2.163	60389
4.4	1.0	27	2.728	2.189	59848
4.5	1.0	27	2.751	2.213	59335
4.6	1.0	27	2.775	2.238	58829
4.7	1.0	27	2.798	2.261	58338
4.8	1.0	30	2.824	2.289	57814
4.9	1.0	30	2.848	2.312	57317
5.0	1.0	30	2.871	2.334	56852
5.1	1.0	30	2.894	2.359	56411
5.2	1.0	30	2.918	2.384	55940
5.3	1.0	30	2.942	2.413	55490
5.4	1.0	30	2.967	2.438	55016
5.5	1.0	30	2.991	2.460	54574
5.6	1.0	30	3.013	2.483	54170
5.7	1.0	30	3.037	2.506	53756
5.8	1.0	30	3.060	2.532	53353
5.9	1.0	30	3.085	2.557	52911
6.0	1.0	30	3.110	2.579	52485
6.1	1.0	30	3.133	2.603	52107
6.2	1.0	30	3.156	2.628	51725
6.3	1.0	30	3.179	2.651	51346
6.4	1.0	30	3.203	2.676	50970
6.5	1.0	30	3.227	2.701	50589
6.6	1.0	30	3.249	2.726	50240
6.7	1.0	30	3.270	2.748	49922
6.8	1.0	30	3.294	2.772	49564
6.9	1.0	30	3.315	2.794	49244
7.0	1.0	30	3.340	2.816	48875
7.1	1.0	30	3.363	2.838	48545
7.2	1.0	30	3.386	2.858	48206
7.3	1.0	30	3.407	2.878	47908
7.4	1.0	30	3.431	2.903	47580
7.5	1.0	30	3.457	2.932	47218
7.6	1.0	30	3.480	2.952	46909
7.7	1.0	30	3.505	2.976	46578
7.8	1.0	30	3.527	2.996	46289
7.9	1.0	30	3.548	3.019	46006
8.0	1.0	30	3.569	3.040	45735
8.1	1.0	30	3.592	3.066	45445
8.2	1.0	30	3.615	3.087	45159
8.3	1.0	30	3.639	3.110	44862
8.4	1.0	30	3.664	3.132	44558
8.5	1.0	30	3.686	3.152	44291
8.6	1.0	30	3.706	3.172	44051
8.7	1.0	30	3.727	3.190	43797
8.8	1.0	30	3.754	3.220	43489
8.9	1.0	30	3.777	3.247	43218
9.0	1.0	46	3.801	3.275	42948
9.1	1.0	46	3.826	3.302	42672
9.2	1.0	46	3.848	3.319	42425
9.3	1.0	46	3.871	3.342	42172
9.4	1.0	46	3.894	3.368	41921
9.5	1.0	46	3.916	3.392	41688
9.6	1.0	46	3.939	3.420	41439
9.7	1.0	46	3.964	3.444	41183
9.8	1.0	46	3.986	3.466	40955
9.9	1.0	46	4.006	3.484	40746
10.0	1.0	46	4.029	3.509	40516

C | Model summaries for Menzerath's law

This chapter of the appendix includes all tables that summarize the model outputs used in chapter 3 for every sequence scenario and used δ. All tables include the estimates of the predictors, the 95% confidence intervals for sequence length, and the AIC difference between full and null model (AIC null - AIC full).

				Confidence intervals		
δ	Intercept	Position ± SE	Sequence Length ± SE	Lower	Upper	AICd
0.1	342.00 ± 5.02	5.0582 ± 0.8390	-2.0191 ± 0.5458	-3.0891	-0.9497	34.47
0.2	339.50 ± 4.99	0.1154 ± 0.1703	0.3583 ± 0.1026	0.1571	0.5593	49.17
0.3	339.57 ± 5.00	0.0135 ± 0.0873	0.2652 ± 0.0532	0.1609	0.3694	76.08
0.4	339.63 ± 5.00	0.0152 ± 0.0613	0.1897 ± 0.0380	0.1153	0.2641	73.71
0.5	339.62 ± 5.01	0.0046 ± 0.0500	0.1674 ± 0.0319	0.1049	0.2298	69.70
0.6	339.77 ± 5.00	0.0041 ± 0.0347	0.0995 ± 0.0227	0.0550	0.1440	44.24
0.7	339.73 ± 5.01	0.0155 ± 0.0307	0.0913 ± 0.0208	0.0505	0.1321	45.92
0.8	339.71 ± 5.01	0.0117 ± 0.0285	0.0874 ± 0.0198	0.0486	0.1262	42.41
0.9	339.87 ± 5.00	-0.0061 ± 0.0228	0.0577 ± 0.0154	0.0274	0.0879	23.89
1.0	340.06 ± 4.98	-0.0071 ± 0.0151	0.0227 ± 0.0097	0.0036	0.0418	5.82
1.1	340.05 ± 4.98	-0.0072 ± 0.0143	0.0231 ± 0.0096	0.0044	0.0419	5.76
1.2	340.10 ± 4.98	-0.0102 ± 0.0110	0.0162 ± 0.0074	0.0016	0.0307	1.78
1.3	340.08 ± 4.98	-0.0094 ± 0.0109	0.0171 ± 0.0075	0.0024	0.0318	2.60
1.4	340.08 ± 4.98	-0.0093 ± 0.0109	0.0173 ± 0.0075	0.0026	0.0320	2.75
1.5	340.08 ± 4.98	-0.0096 ± 0.0107	0.0168 ± 0.0076	0.0018	0.0317	1.74
1.6	340.07 ± 4.98	-0.0094 ± 0.0107	0.0172 ± 0.0076	0.0022	0.0321	2.11
1.7	340.06 ± 4.98	-0.0097 ± 0.0107	0.0179 ± 0.0076	0.0030	0.0328	2.68
1.8	340.06 ± 4.99	-0.0111 ± 0.0106	0.0183 ± 0.0076	0.0033	0.0333	2.41
1.9	340.05 ± 4.99	-0.0109 ± 0.0106	0.0182 ± 0.0076	0.0032	0.0332	2.42
2.0	340.05 ± 4.99	-0.0108 ± 0.0106	0.0185 ± 0.0076	0.0035	0.0334	2.64
2.1	340.04 ± 4.99	-0.0099 ± 0.0106	0.0185 ± 0.0077	0.0035	0.0335	2.85
2.2	340.03 ± 4.99	-0.0098 ± 0.0106	0.0187 ± 0.0077	0.0037	0.0337	3.02
2.3	340.02 ± 4.99	-0.0095 ± 0.0106	0.0190 ± 0.0076	0.0040	0.0339	3.38
2.4	340.01 ± 4.99	-0.0098 ± 0.0105	0.0196 ± 0.0076	0.0047	0.0346	3.96
2.5	340.00 ± 4.99	-0.0107 ± 0.0105	0.0205 ± 0.0077	0.0055	0.0355	4.42
2.6	339.99 ± 4.99	-0.0107 ± 0.0105	0.0208 ± 0.0077	0.0058	0.0358	4.76
2.7	339.80 ± 5.01	-0.0120 ± 0.0100	0.0344 ± 0.0077	0.0194	0.0495	21.28
2.8	339.77 ± 5.02	-0.0122 ± 0.0099	0.0360 ± 0.0077	0.0210	0.0511	23.59
2.9	339.76 ± 5.02	-0.0115 ± 0.0099	0.0362 ± 0.0077	0.0212	0.0513	24.28
3.0	339.74 ± 5.02	-0.0115 ± 0.0099	0.0364 ± 0.0077	0.0214	0.0515	24.65
3.1	339.74 ± 5.02	-0.0117 ± 0.0098	0.0359 ± 0.0078	0.0206	0.0512	21.99
3.2	339.73 ± 5.02	-0.0115 ± 0.0098	0.0363 ± 0.0078	0.0210	0.0515	22.61
3.3	339.72 ± 5.02	-0.0116 ± 0.0098	0.0362 ± 0.0078	0.0209	0.0515	22.41
3.4	339.72 ± 5.02	-0.0115 ± 0.0098	0.0357 ± 0.0078	0.0204	0.0510	21.70
3.5	339.72 ± 5.02	-0.0115 ± 0.0098	0.0354 ± 0.0078	0.0201	0.0507	21.21
3.6	339.71 ± 5.02	-0.0106 ± 0.0096	0.0350 ± 0.0079	0.0194	0.0505	19.46
3.7	339.69 ± 5.02	-0.0108 ± 0.0096	0.0359 ± 0.0080	0.0202	0.0516	19.87
3.8	339.67 ± 5.03	-0.0111 ± 0.0096	0.0370 ± 0.0080	0.0213	0.0526	21.29
3.9	339.67 ± 5.03	-0.0112 ± 0.0095	0.0365 ± 0.0080	0.0208	0.0521	20.62
4.0	339.65 ± 5.03	-0.0112 ± 0.0095	0.0373 ± 0.0080	0.0217	0.0529	21.92
4.1	339.64 ± 5.03	-0.0108 ± 0.0095	0.0374 ± 0.0080	0.0217	0.0530	22.16
4.2	339.63 ± 5.03	-0.0107 ± 0.0095	0.0371 ± 0.0080	0.0215	0.0527	21.83
4.3	339.61 ± 5.03	-0.0118 ± 0.0095	0.0387 ± 0.0080	0.0230	0.0543	23.40
4.4	339.50 ± 5.05	-0.0126 ± 0.0087	0.0436 ± 0.0080	0.0278	0.0593	28.60

Table 15: LMM summary results for values of δ=[0.1, 4.4] - initial sequences where no shufflings were performed

				Confidence intervals		
δ	Intercept	Position ± SE	Sequence Length ± SE	Lower	Upper	AICd
4.5	339.46 ± 5.05	-0.0127 ± 0.0087	0.0438 ± 0.0080	0.0281	0.0595	29.02
4.6	339.45 ± 5.05	-0.0126 ± 0.0087	0.0433 ± 0.0080	0.0276	0.0589	28.52
4.7	339.45 ± 5.05	-0.0125 ± 0.0087	0.0433 ± 0.0080	0.0276	0.0589	28.63
4.8	339.44 ± 5.05	-0.0122 ± 0.0087	0.0430 ± 0.0080	0.0273	0.0586	28.21
4.9	339.43 ± 5.06	-0.0121 ± 0.0087	0.0430 ± 0.0080	0.0274	0.0586	28.53
5.0	339.43 ± 5.06	-0.0120 ± 0.0087	0.0427 ± 0.0080	0.0271	0.0583	28.13
5.1	339.40 ± 5.06	-0.0121 ± 0.0087	0.0434 ± 0.0079	0.0278	0.0590	29.39
5.2	339.38 ± 5.06	-0.0116 ± 0.0086	0.0437 ± 0.0079	0.0280	0.0592	29.90
5.3	339.42 ± 5.06	-0.0116 ± 0.0086	0.0411 ± 0.0078	0.0257	0.0564	26.63
5.4	339.41 ± 5.06	-0.0110 ± 0.0086	0.0407 ± 0.0078	0.0253	0.0560	26.63
5.5	339.43 ± 5.06	-0.0111 ± 0.0086	0.0393 ± 0.0077	0.0241	0.0545	25.08
5.6	339.41 ± 5.06	-0.0112 ± 0.0085	0.0399 ± 0.0077	0.0247	0.0551	25.99
5.7	339.39 ± 5.06	-0.0112 ± 0.0085	0.0406 ± 0.0077	0.0254	0.0558	27.13
5.8	339.38 ± 5.06	-0.0112 ± 0.0085	0.0409 ± 0.0077	0.0257	0.0560	27.55
5.9	339.36 ± 5.06	-0.0112 ± 0.0085	0.0416 ± 0.0077	0.0264	0.0567	28.77
6.0	339.35 ± 5.06	-0.0111 ± 0.0085	0.0418 ± 0.0077	0.0266	0.0570	29.27
6.1	339.34 ± 5.06	-0.0110 ± 0.0085	0.0419 ± 0.0077	0.0268	0.0571	29.58
6.2	339.32 ± 5.07	-0.0111 ± 0.0085	0.0428 ± 0.0077	0.0276	0.0579	31.15
6.3	339.33 ± 5.06	-0.0100 ± 0.0085	0.0412 ± 0.0077	0.0261	0.0562	29.17
6.4	339.30 ± 5.07	-0.0106 ± 0.0085	0.0425 ± 0.0077	0.0274	0.0576	31.25
6.5	339.28 ± 5.07	-0.0110 ± 0.0085	0.0434 ± 0.0077	0.0283	0.0584	32.60
6.6	339.26 ± 5.07	-0.0111 ± 0.0085	0.0440 ± 0.0077	0.0289	0.0591	33.80
6.7	339.25 ± 5.07	-0.0110 ± 0.0085	0.0443 ± 0.0077	0.0292	0.0593	34.31
6.8	339.25 ± 5.07	-0.0111 ± 0.0085	0.0437 ± 0.0077	0.0287	0.0588	33.37
6.9	339.39 ± 5.06	-0.0134 ± 0.0081	0.0388 ± 0.0073	0.0244	0.0531	26.48
7.0	339.37 ± 5.06	-0.0132 ± 0.0081	0.0392 ± 0.0073	0.0248	0.0535	27.26
7.1	339.35 ± 5.06	-0.0132 ± 0.0081	0.0400 ± 0.0073	0.0256	0.0543	28.62
7.2	339.33 ± 5.06	-0.0126 ± 0.0081	0.0403 ± 0.0073	0.0260	0.0546	29.65
7.3	339.32 ± 5.06	-0.0126 ± 0.0081	0.0406 ± 0.0073	0.0262	0.0549	30.17
7.4	339.30 ± 5.06	-0.0124 ± 0.0081	0.0408 ± 0.0073	0.0264	0.0551	30.68
7.5	339.29 ± 5.06	-0.0125 ± 0.0081	0.0411 ± 0.0073	0.0268	0.0554	31.19
7.6	339.28 ± 5.07	-0.0125 ± 0.0081	0.0412 ± 0.0073	0.0269	0.0555	31.35
7.7	339.27 ± 5.07	-0.0127 ± 0.0081	0.0418 ± 0.0073	0.0274	0.0561	32.33
7.8	339.26 ± 5.07	-0.0126 ± 0.0081	0.0420 ± 0.0073	0.0277	0.0563	32.85
7.9	339.19 ± 5.07	-0.0126 ± 0.0081	0.0440 ± 0.0073	0.0297	0.0582	37.16
8.0	339.18 ± 5.07	-0.0125 ± 0.0081	0.0440 ± 0.0073	0.0297	0.0582	37.24
8.1	339.15 ± 5.08	-0.0123 ± 0.0080	0.0448 ± 0.0073	0.0305	0.0590	38.92
8.2	339.14 ± 5.08	-0.0125 ± 0.0080	0.0451 ± 0.0073	0.0308	0.0593	39.47
8.3	339.14 ± 5.08	-0.0125 ± 0.0080	0.0451 ± 0.0073	0.0308	0.0593	39.50
8.4	339.14 ± 5.08	-0.0124 ± 0.0080	0.0451 ± 0.0073	0.0308	0.0593	39.57
8.5	339.10 ± 5.08	-0.0123 ± 0.0080	0.0463 ± 0.0073	0.0320	0.0605	42.33
8.6	339.09 ± 5.08	-0.0121 ± 0.0080	0.0465 ± 0.0073	0.0322	0.0607	42.85
8.7	339.07 ± 5.08	-0.0121 ± 0.0080	0.0471 ± 0.0073	0.0328	0.0613	44.15
8.8	339.06 ± 5.08	-0.0123 ± 0.0080	0.0471 ± 0.0073	0.0328	0.0613	44.10
8.9	339.05 ± 5.09	-0.0123 ± 0.0080	0.0472 ± 0.0073	0.0330	0.0615	44.44
9.0	339.04 ± 5.09	-0.0123 ± 0.0080	0.0476 ± 0.0073	0.0333	0.0618	45.25
9.1	339.02 ± 5.09	-0.0123 ± 0.0080	0.0479 ± 0.0072	0.0336	0.0621	45.98
9.2	339.00 ± 5.09	-0.0121 ± 0.0080	0.0483 ± 0.0073	0.0341	0.0625	47.06
9.3	339.00 ± 5.09	-0.0121 ± 0.0080	0.0477 ± 0.0072	0.0335	0.0619	46.22
9.4	339.00 ± 5.09	-0.0121 ± 0.0080	0.0478 ± 0.0072	0.0336	0.0619	46.33
9.5	338.98 ± 5.09	-0.0120 ± 0.0080	0.0484 ± 0.0072	0.0342	0.0625	47.80
9.6	338.96 ± 5.09	-0.0120 ± 0.0080	0.0485 ± 0.0072	0.0344	0.0627	48.28
9.7	338.94 ± 5.10	-0.0118 ± 0.0080	0.0485 ± 0.0072	0.0344	0.0626	48.89
9.8	338.93 ± 5.10	-0.0111 ± 0.0080	0.0484 ± 0.0072	0.0343	0.0625	49.30
9.9	338.92 ± 5.10	-0.0108 ± 0.0080	0.0481 ± 0.0072	0.0340	0.0622	48.92
10.0	338.91 ± 5.10	-0.0109 ± 0.0080	0.0483 ± 0.0072	0.0342	0.0624	49.35

Table 16: LMM summary results for values of δ=[0.1, 4.4] - initial sequences where to shufflings were performed

				Confidence intervals		
δ	Intercept	Position ± SE	Sequence Length ± SE	Lower	Upper	AICd
0.1	341.66 ± 5.11	-0.5633 ± 1.3602	-0.9578 ± 0.9653	-2.8497	0.9342	-0.5550
0.2	340.92 ± 5.07	0.0199 ± 1.0851	-0.4531 ± 0.7685	-1.9593	1.0532	-3.3368
0.3	341.26 ± 5.06	0.0831 ± 0.9535	-0.6318 ± 0.6741	-1.9530	0.6894	-2.4584
0.4	340.92 ± 5.05	-0.1062 ± 0.8577	-0.3628 ± 0.6069	-1.5523	0.8268	-3.0470
0.5	341.02 ± 5.05	-0.1676 ± 0.7883	-0.3760 ± 0.5589	-1.4714	0.7195	-2.6088
0.6	340.52 ± 5.04	-0.2313 ± 0.7286	-0.0900 ± 0.5159	-1.1011	0.9211	-3.5823
0.7	340.72 ± 5.04	-0.1269 ± 0.6860	-0.2007 ± 0.4864	-1.1540	0.7527	-3.3795
0.8	340.54 ± 5.04	-0.0289 ± 0.6489	-0.1382 ± 0.4595	-1.0388	0.7625	-3.7781
0.9	340.63 ± 5.03	0.0776 ± 0.6165	-0.1996 ± 0.4365	-1.0552	0.6560	-3.7137
1.0	340.15 ± 5.03	-0.0540 ± 0.5896	0.0381 ± 0.4173	-0.7798	0.8561	-3.9902
1.1	339.93 ± 5.03	-0.0862 ± 0.5675	0.1317 ± 0.4013	-0.6547	0.9182	-3.8793
1.2	340.01 ± 5.03	-0.1205 ± 0.5477	0.1105 ± 0.3870	-0.6481	0.8691	-3.9178
1.3	339.93 ± 5.03	0.1207 ± 0.5302	0.0620 ± 0.3748	-0.6726	0.7965	-3.7348
1.4	339.61 ± 5.03	-0.0322 ± 0.5141	0.2238 ± 0.3631	-0.4878	0.9355	-3.3397
1.5	339.64 ± 5.03	-0.1899 ± 0.5008	0.2594 ± 0.3541	-0.4347	0.9535	-3.4252
1.6	339.30 ± 5.03	-0.1769 ± 0.4882	0.3698 ± 0.3453	-0.3069	1.0466	-2.5411
1.7	339.13 ± 5.03	-0.1304 ± 0.4758	0.4038 ± 0.3366	-0.2558	1.0635	-1.9019
1.8	339.15 ± 5.03	-0.1625 ± 0.4652	0.4005 ± 0.3292	-0.2447	1.0458	-1.9995
1.9	339.10 ± 5.02	-0.0333 ± 0.4548	0.3684 ± 0.3218	-0.2623	0.9992	-1.6083
2.0	339.28 ± 5.02	-0.0409 ± 0.4453	0.3072 ± 0.3152	-0.3106	0.9251	-2.3399
2.1	339.38 ± 5.02	-0.0405 ± 0.4352	0.2706 ± 0.3080	-0.3330	0.8742	-2.6717
2.2	339.45 ± 5.02	-0.0872 ± 0.4261	0.2599 ± 0.3016	-0.3311	0.8511	-2.9307
2.3	339.32 ± 5.02	-0.1876 ± 0.4187	0.3282 ± 0.2965	-0.2529	0.9094	-2.5521
2.4	338.95 ± 5.02	-0.2992 ± 0.4098	0.4745 ± 0.2898	-0.0936	1.0425	-0.9550
2.5	338.98 ± 5.02	-0.3425 ± 0.4025	0.4757 ± 0.2845	-0.0819	1.0332	-0.9833
2.6	338.82 ± 5.02	-0.3572 ± 0.3948	0.5205 ± 0.2789	-0.0260	1.0671	-0.1683
2.7	338.75 ± 5.02	-0.2353 ± 0.3873	0.4892 ± 0.2732	-0.0462	1.0246	0.0875
2.8	338.66 ± 5.02	-0.2876 ± 0.3814	0.5284 ± 0.2691	0.0010	1.0558	0.6736
2.9	338.91 ± 5.02	-0.2877 ± 0.3755	0.4557 ± 0.2650	-0.0636	0.9751	-0.6304
3.0	338.89 ± 5.02	-0.2292 ± 0.3698	0.4354 ± 0.2609	-0.0759	0.9467	-0.5780
3.1	339.00 ± 5.02	-0.2064 ± 0.3637	0.3923 ± 0.2566	-0.1105	0.8952	-1.1246
3.2	338.73 ± 5.02	-0.1810 ± 0.3586	0.4506 ± 0.2530	-0.0453	0.9464	0.3247
3.3	338.94 ± 5.02	-0.2123 ± 0.3535	0.4031 ± 0.2493	-0.0855	0.8918	-0.7864
3.4	338.75 ± 5.02	-0.2621 ± 0.3492	0.4678 ± 0.2464	-0.0150	0.9507	0.3181
3.5	338.81 ± 5.02	-0.3045 ± 0.3452	0.4639 ± 0.2439	-0.0141	0.9419	0.0499
3.6	338.87 ± 5.02	-0.3656 ± 0.3402	0.4672 ± 0.2402	-0.0035	0.9379	-0.0314
3.7	338.62 ± 5.02	-0.4542 ± 0.3354	0.5602 ± 0.2368	0.0962	1.0242	1.8060
3.8	338.72 ± 5.02	-0.4553 ± 0.3306	0.5319 ± 0.2333	0.0747	0.9892	1.3130
3.9	338.79 ± 5.02	-0.5277 ± 0.3263	0.5386 ± 0.2303	0.0873	0.9900	1.4741
4.0	338.54 ± 5.02	-0.5971 ± 0.3222	0.6198 ± 0.2274	0.1742	1.0655	3.4440
4.1	338.70 ± 5.02	-0.5814 ± 0.3183	0.5726 ± 0.2248	0.1321	1.0131	2.4916
4.2	338.69 ± 5.02	-0.4894 ± 0.3144	0.5380 ± 0.2220	0.1029	0.9731	1.9257
4.3	338.98 ± 5.02	-0.4513 ± 0.3112	0.4527 ± 0.2198	0.0219	0.8835	0.2420
4.4	339.14 ± 5.02	-0.4464 ± 0.3074	0.4120 ± 0.2171	-0.0135	0.8375	-0.3741
4.5	338.76 ± 5.02	-0.4416 ± 0.3040	0.4934 ± 0.2147	0.0726	0.9143	1.3418
4.6	338.75 ± 5.02	-0.4851 ± 0.3006	0.5103 ± 0.2123	0.0941	0.9265	1.7901
4.7	338.72 ± 5.02	-0.4710 ± 0.2976	0.5076 ± 0.2103	0.0954	0.9198	1.8555
4.8	338.67 ± 5.02	-0.3758 ± 0.2940	0.4780 ± 0.2077	0.0709	0.8851	1.5420
4.9	338.45 ± 5.02	-0.3332 ± 0.2913	0.5054 ± 0.2060	0.1017	0.9091	2.7202
5.0	338.55 ± 5.02	-0.2536 ± 0.2886	0.4494 ± 0.2042	0.0491	0.8497	1.7564

Table 17: LMM summary results for values of δ=[0.1, 5.0] - sequences produced after shuffling the order of silences between the whistles

				Confidence intervals		
δ	Intercept	Position ± SE	Sequence Length ± SE	Lower	Upper	AICd
5.1	338.41 ± 5.02	-0.2753 ± 0.2855	0.4837 ± 0.2019	0.0879	0.8795	2.8008
5.2	338.37 ± 5.02	-0.2735 ± 0.2825	0.4880 ± 0.1998	0.0964	0.8796	3.1160
5.3	338.51 ± 5.02	-0.2213 ± 0.2790	0.4355 ± 0.1971	0.0492	0.8218	2.0704
5.4	338.63 ± 5.02	-0.2317 ± 0.2761	0.4111 ± 0.1951	0.0287	0.7934	1.2900
5.5	338.65 ± 5.02	-0.2818 ± 0.2738	0.4260 ± 0.1936	0.0466	0.8054	1.3966
5.6	338.84 ± 5.02	-0.2913 ± 0.2713	0.3884 ± 0.1919	0.0123	0.7645	0.3510
5.7	339.02 ± 5.02	-0.3044 ± 0.2688	0.3557 ± 0.1902	-0.0171	0.7284	-0.4314
5.8	339.08 ± 5.02	-0.2292 ± 0.2660	0.3125 ± 0.1880	-0.0560	0.6810	-1.0395
5.9	339.16 ± 5.02	-0.1508 ± 0.2635	0.2630 ± 0.1864	-0.1023	0.6283	-1.6474
6.0	339.08 ± 5.02	-0.1115 ± 0.2614	0.2590 ± 0.1850	-0.1037	0.6217	-1.4105
6.1	339.16 ± 5.02	-0.0826 ± 0.2591	0.2313 ± 0.1834	-0.1282	0.5909	-1.7569
6.2	338.96 ± 5.02	-0.1763 ± 0.2564	0.3047 ± 0.1814	-0.0508	0.6603	-0.6802
6.3	339.07 ± 5.02	-0.1572 ± 0.2542	0.2756 ± 0.1799	-0.0770	0.6283	-1.2221
6.4	339.01 ± 5.02	-0.1235 ± 0.2517	0.2713 ± 0.1781	-0.0778	0.6204	-0.9938
6.5	339.06 ± 5.02	-0.1372 ± 0.2493	0.2662 ± 0.1763	-0.0794	0.6118	-1.1870
6.6	338.80 ± 5.02	-0.1911 ± 0.2469	0.3328 ± 0.1745	-0.0092	0.6748	0.3014
6.7	338.77 ± 5.02	-0.1926 ± 0.2449	0.3375 ± 0.1731	-0.0018	0.6768	0.5051
6.8	338.80 ± 5.02	-0.2351 ± 0.2428	0.3473 ± 0.1717	0.0109	0.6838	0.5238
6.9	338.49 ± 5.02	-0.2834 ± 0.2410	0.4203 ± 0.1704	0.0864	0.7543	2.7354
7.0	338.44 ± 5.02	-0.2720 ± 0.2391	0.4234 ± 0.1692	0.0919	0.7550	3.0643
7.1	338.49 ± 5.02	-0.3134 ± 0.2374	0.4290 ± 0.1680	0.0996	0.7583	2.9836
7.2	338.46 ± 5.02	-0.3424 ± 0.2358	0.4442 ± 0.1670	0.1169	0.7715	3.4380
7.3	338.50 ± 5.02	-0.2908 ± 0.2342	0.4140 ± 0.1659	0.0888	0.7391	2.7664
7.4	338.44 ± 5.02	-0.3367 ± 0.2322	0.4419 ± 0.1645	0.1196	0.7643	3.6178
7.5	338.67 ± 5.02	-0.3721 ± 0.2299	0.4157 ± 0.1628	0.0967	0.7348	2.5903
7.6	338.77 ± 5.02	-0.3910 ± 0.2284	0.4046 ± 0.1619	0.0874	0.7219	2.2560
7.7	338.76 ± 5.02	-0.4017 ± 0.2267	0.4097 ± 0.1608	0.0947	0.7248	2.4978
7.8	338.78 ± 5.02	-0.4097 ± 0.2252	0.4082 ± 0.1598	0.0951	0.7214	2.5292
7.9	338.84 ± 5.02	-0.3949 ± 0.2234	0.3909 ± 0.1584	0.0804	0.7014	2.0892
8.0	338.78 ± 5.02	-0.3589 ± 0.2220	0.3839 ± 0.1575	0.0753	0.6927	1.9632
8.1	338.73 ± 5.02	-0.3383 ± 0.2200	0.3819 ± 0.1560	0.0763	0.6877	2.0687
8.2	338.76 ± 5.02	-0.3410 ± 0.2186	0.3772 ± 0.1550	0.0734	0.6810	1.9706
8.3	338.69 ± 5.02	-0.3348 ± 0.2170	0.3836 ± 0.1539	0.0820	0.6852	2.3039
8.4	338.75 ± 5.02	-0.3434 ± 0.2155	0.3774 ± 0.1529	0.0777	0.6771	2.1353
8.5	338.64 ± 5.02	-0.3505 ± 0.2143	0.3965 ± 0.1522	0.0981	0.6948	2.8618
8.6	338.70 ± 5.02	-0.3614 ± 0.2130	0.3903 ± 0.1513	0.0938	0.6868	2.6862
8.7	338.66 ± 5.02	-0.3552 ± 0.2118	0.3923 ± 0.1506	0.0971	0.6874	2.8349
8.8	338.81 ± 5.02	-0.3727 ± 0.2098	0.3753 ± 0.1491	0.0831	0.6675	2.3376
8.9	338.83 ± 5.02	-0.3911 ± 0.2081	0.3782 ± 0.1479	0.0883	0.6680	2.5529
9.0	338.46 ± 5.02	-0.4917 ± 0.2060	0.4759 ± 0.1461	0.1896	0.7622	6.6289
9.1	338.47 ± 5.02	-0.5347 ± 0.2043	0.4900 ± 0.1449	0.2060	0.7740	7.5427
9.2	338.42 ± 5.02	-0.5032 ± 0.2034	0.4834 ± 0.1443	0.2005	0.7663	7.2398
9.3	338.26 ± 5.01	-0.5116 ± 0.2021	0.5096 ± 0.1435	0.2284	0.7908	8.6167
9.4	338.31 ± 5.01	-0.5095 ± 0.2005	0.4978 ± 0.1424	0.2187	0.7769	8.2335
9.5	338.22 ± 5.01	-0.5105 ± 0.1991	0.5109 ± 0.1414	0.2338	0.7880	9.0589
9.6	338.17 ± 5.01	-0.4143 ± 0.1974	0.4743 ± 0.1401	0.1997	0.7489	7.6233
9.7	338.07 ± 5.01	-0.4048 ± 0.1959	0.4837 ± 0.1390	0.2113	0.7561	8.4083
9.8	338.00 ± 5.01	-0.4304 ± 0.1947	0.5024 ± 0.1381	0.2317	0.7732	9.4760
9.9	337.92 ± 5.01	-0.4286 ± 0.1938	0.5116 ± 0.1375	0.2420	0.7812	10.1661
10.0	337.98 ± 5.01	-0.4279 ± 0.1924	0.5007 ± 0.1366	0.2330	0.7685	9.6887

Table 18: LMM summary results for values values of δ=[5.1, 10.0] - sequences produced after shuffling the order of silences between the whistles

				Confidence intervals		
δ	Intercept	Position ± SE	Sequence Length ± SE	Lower	Upper	AICd
0.1	340.59 ± 5.01	0.3595 ± 0.8395	-0.3078 ± 0.5183	-1.3235	0.7080	-3.63
0.2	340.35 ± 4.97	0.0025 ± 0.1704	-0.0559 ± 0.0991	-0.2502	0.1384	-2.83
0.3	340.30 ± 4.97	0.1254 ± 0.0874	-0.0708 ± 0.0507	-0.1702	0.0286	-1.84
0.4	340.26 ± 4.97	0.0671 ± 0.0614	-0.0372 ± 0.0356	-0.1070	0.0326	-2.76
0.5	340.25 ± 4.97	0.0383 ± 0.0500	-0.0223 ± 0.0290	-0.0792	0.0346	-3.36
0.6	340.23 ± 4.97	0.0054 ± 0.0348	-0.0046 ± 0.0202	-0.0442	0.0349	-3.93
0.7	340.26 ± 4.97	-0.0054 ± 0.0307	-0.0022 ± 0.0178	-0.0371	0.0328	-3.67
0.8	340.21 ± 4.97	-0.0017 ± 0.0285	0.0002 ± 0.0166	-0.0323	0.0327	-3.99
0.9	340.20 ± 4.97	-0.0036 ± 0.0229	0.0016 ± 0.0133	-0.0244	0.0276	-3.97
1.0	340.17 ± 4.97	0.0069 ± 0.0151	-0.0015 ± 0.0087	-0.0186	0.0156	-3.59
1.1	340.19 ± 4.97	0.0067 ± 0.0143	-0.0025 ± 0.0083	-0.0188	0.0138	-3.73
1.2	340.20 ± 4.97	0.0164 ± 0.0110	-0.0078 ± 0.0064	-0.0203	0.0046	-1.75
1.3	340.21 ± 4.97	0.0176 ± 0.0109	-0.0086 ± 0.0063	-0.0210	0.0038	-1.39
1.4	340.21 ± 4.97	0.0169 ± 0.0109	-0.0083 ± 0.0063	-0.0207	0.0041	-1.60
1.5	340.22 ± 4.97	0.0152 ± 0.0108	-0.0077 ± 0.0062	-0.0200	0.0045	-2.01
1.6	340.22 ± 4.97	0.0152 ± 0.0107	-0.0079 ± 0.0062	-0.0201	0.0044	-2.00
1.7	340.23 ± 4.97	0.0160 ± 0.0107	-0.0084 ± 0.0062	-0.0206	0.0038	-1.75
1.8	340.21 ± 4.97	0.0166 ± 0.0106	-0.0083 ± 0.0062	-0.0204	0.0038	-1.55
1.9	340.21 ± 4.97	0.0164 ± 0.0106	-0.0082 ± 0.0062	-0.0203	0.0039	-1.61
2.0	340.21 ± 4.97	0.0163 ± 0.0106	-0.0081 ± 0.0062	-0.0202	0.0039	-1.65
2.1	340.21 ± 4.97	0.0163 ± 0.0106	-0.0080 ± 0.0061	-0.0201	0.0040	-1.62
2.2	340.21 ± 4.97	0.0172 ± 0.0106	-0.0084 ± 0.0061	-0.0205	0.0036	-1.33
2.3	340.21 ± 4.97	0.0174 ± 0.0106	-0.0085 ± 0.0061	-0.0205	0.0035	-1.28
2.4	340.21 ± 4.97	0.0177 ± 0.0105	-0.0087 ± 0.0061	-0.0207	0.0033	-1.16
2.5	340.21 ± 4.97	0.0180 ± 0.0105	-0.0089 ± 0.0061	-0.0209	0.0031	-1.06
2.6	340.21 ± 4.97	0.0176 ± 0.0105	-0.0085 ± 0.0061	-0.0205	0.0034	-1.21
2.7	340.22 ± 4.97	0.0155 ± 0.0100	-0.0079 ± 0.0058	-0.0193	0.0035	-1.59
2.8	340.22 ± 4.97	0.0156 ± 0.0099	-0.0081 ± 0.0058	-0.0194	0.0032	-1.53
2.9	340.22 ± 4.97	0.0152 ± 0.0099	-0.0080 ± 0.0058	-0.0192	0.0033	-1.63
3.0	340.23 ± 4.97	0.0160 ± 0.0099	-0.0084 ± 0.0058	-0.0197	0.0029	-1.38
3.1	340.24 ± 4.97	0.0136 ± 0.0098	-0.0075 ± 0.0057	-0.0187	0.0036	-2.00
3.2	340.23 ± 4.97	0.0137 ± 0.0098	-0.0075 ± 0.0057	-0.0186	0.0037	-1.99
3.3	340.22 ± 4.97	0.0156 ± 0.0098	-0.0081 ± 0.0057	-0.0192	0.0031	-1.45
3.4	340.22 ± 4.97	0.0158 ± 0.0098	-0.0082 ± 0.0057	-0.0193	0.0029	-1.38
3.5	340.22 ± 4.97	0.0159 ± 0.0098	-0.0083 ± 0.0057	-0.0194	0.0028	-1.33
3.6	340.24 ± 4.97	0.0140 ± 0.0096	-0.0077 ± 0.0056	-0.0186	0.0033	-1.82
3.7	340.24 ± 4.97	0.0133 ± 0.0096	-0.0075 ± 0.0056	-0.0184	0.0034	-1.96
3.8	340.24 ± 4.97	0.0134 ± 0.0096	-0.0075 ± 0.0056	-0.0184	0.0034	-1.93
3.9	340.24 ± 4.97	0.0137 ± 0.0095	-0.0077 ± 0.0055	-0.0185	0.0032	-1.85
4.0	340.24 ± 4.97	0.0134 ± 0.0095	-0.0075 ± 0.0055	-0.0184	0.0034	-1.93
4.1	340.25 ± 4.97	0.0136 ± 0.0095	-0.0077 ± 0.0055	-0.0186	0.0032	-1.85
4.2	340.24 ± 4.97	0.0128 ± 0.0095	-0.0072 ± 0.0055	-0.0180	0.0037	-2.12
4.3	340.23 ± 4.97	0.0140 ± 0.0095	-0.0075 ± 0.0055	-0.0184	0.0033	-1.77
4.4	340.27 ± 4.97	0.0114 ± 0.0087	-0.0069 ± 0.0051	-0.0169	0.0030	-2.06
4.5	340.27 ± 4.97	0.0129 ± 0.0087	-0.0076 ± 0.0051	-0.0175	0.0023	-1.60
4.6	340.26 ± 4.97	0.0137 ± 0.0087	-0.0078 ± 0.0051	-0.0177	0.0021	-1.36
4.7	340.26 ± 4.97	0.0136 ± 0.0087	-0.0078 ± 0.0051	-0.0177	0.0021	-1.39
4.8	340.26 ± 4.97	0.0135 ± 0.0087	-0.0077 ± 0.0051	-0.0176	0.0022	-1.44
4.9	340.26 ± 4.97	0.0134 ± 0.0087	-0.0078 ± 0.0050	-0.0177	0.0021	-1.45
5.0	340.26 ± 4.97	0.0133 ± 0.0087	-0.0077 ± 0.0050	-0.0176	0.0022	-1.48

Table 19: LMM summary results for values of δ=[0.1, 5.0] - sequences produced after shuffling the order of whistles' duration

				Confidence intervals		
δ	Intercept	Position ± SE	Sequence Length ± SE	Lower	Upper	AICd
5.1	340.26 ± 4.97	0.0130 ± 0.0087	-0.0076 ± 0.0050	-0.0174	0.0023	-1.56
5.2	340.26 ± 4.97	0.0128 ± 0.0087	-0.0074 ± 0.0050	-0.0173	0.0024	-1.64
5.3	340.26 ± 4.97	0.0133 ± 0.0086	-0.0077 ± 0.0050	-0.0175	0.0022	-1.47
5.4	340.27 ± 4.97	0.0135 ± 0.0086	-0.0079 ± 0.0050	-0.0177	0.0019	-1.32
5.5	340.27 ± 4.97	0.0137 ± 0.0086	-0.0080 ± 0.0050	-0.0177	0.0018	-1.24
5.6	340.27 ± 4.97	0.0139 ± 0.0086	-0.0081 ± 0.0050	-0.0178	0.0017	-1.15
5.7	340.27 ± 4.97	0.0138 ± 0.0086	-0.0080 ± 0.0050	-0.0177	0.0018	-1.23
5.8	340.27 ± 4.97	0.0138 ± 0.0086	-0.0080 ± 0.0050	-0.0178	0.0018	-1.21
5.9	340.27 ± 4.97	0.0139 ± 0.0086	-0.0080 ± 0.0050	-0.0178	0.0018	-1.19
6.0	340.27 ± 4.97	0.0138 ± 0.0086	-0.0080 ± 0.0050	-0.0177	0.0018	-1.21
6.1	340.27 ± 4.97	0.0141 ± 0.0085	-0.0081 ± 0.0050	-0.0179	0.0017	-1.11
6.2	340.27 ± 4.97	0.0141 ± 0.0085	-0.0081 ± 0.0050	-0.0179	0.0016	-1.08
6.3	340.27 ± 4.97	0.0142 ± 0.0085	-0.0082 ± 0.0050	-0.0180	0.0015	-1.05
6.4	340.27 ± 4.97	0.0143 ± 0.0085	-0.0083 ± 0.0050	-0.0181	0.0014	-0.99
6.5	340.27 ± 4.97	0.0143 ± 0.0085	-0.0083 ± 0.0050	-0.0181	0.0014	-0.99
6.6	340.28 ± 4.97	0.0144 ± 0.0085	-0.0084 ± 0.0050	-0.0181	0.0014	-0.93
6.7	340.28 ± 4.97	0.0143 ± 0.0085	-0.0084 ± 0.0050	-0.0181	0.0014	-0.96
6.8	340.28 ± 4.97	0.0145 ± 0.0085	-0.0085 ± 0.0050	-0.0182	0.0012	-0.85
6.9	340.23 ± 4.97	0.0168 ± 0.0081	-0.0087 ± 0.0047	-0.0180	0.0005	0.30
7.0	340.24 ± 4.97	0.0167 ± 0.0081	-0.0087 ± 0.0047	-0.0180	0.0005	0.25
7.1	340.24 ± 4.97	0.0168 ± 0.0081	-0.0088 ± 0.0047	-0.0180	0.0005	0.32
7.2	340.23 ± 4.97	0.0166 ± 0.0081	-0.0087 ± 0.0047	-0.0179	0.0006	0.23
7.3	340.23 ± 4.97	0.0165 ± 0.0081	-0.0086 ± 0.0047	-0.0179	0.0006	0.15
7.4	340.23 ± 4.97	0.0159 ± 0.0081	-0.0083 ± 0.0047	-0.0176	0.0009	-0.09
7.5	340.23 ± 4.97	0.0154 ± 0.0081	-0.0081 ± 0.0047	-0.0173	0.0011	-0.32
7.6	340.23 ± 4.97	0.0154 ± 0.0081	-0.0081 ± 0.0047	-0.0173	0.0012	-0.34
7.7	340.23 ± 4.97	0.0154 ± 0.0081	-0.0081 ± 0.0047	-0.0173	0.0012	-0.33
7.8	340.23 ± 4.97	0.0154 ± 0.0081	-0.0080 ± 0.0047	-0.0173	0.0012	-0.35
7.9	340.23 ± 4.97	0.0150 ± 0.0081	-0.0078 ± 0.0047	-0.0170	0.0014	-0.50
8.0	340.23 ± 4.97	0.0151 ± 0.0081	-0.0078 ± 0.0047	-0.0171	0.0014	-0.47
8.1	340.23 ± 4.97	0.0150 ± 0.0081	-0.0078 ± 0.0047	-0.0170	0.0014	-0.51
8.2	340.23 ± 4.97	0.0150 ± 0.0081	-0.0079 ± 0.0047	-0.0171	0.0013	-0.50
8.3	340.24 ± 4.97	0.0151 ± 0.0081	-0.0080 ± 0.0047	-0.0172	0.0012	-0.43
8.4	340.24 ± 4.97	0.0152 ± 0.0080	-0.0081 ± 0.0047	-0.0173	0.0011	-0.37
8.5	340.24 ± 4.97	0.0154 ± 0.0080	-0.0081 ± 0.0047	-0.0173	0.0011	-0.31
8.6	340.24 ± 4.97	0.0154 ± 0.0080	-0.0081 ± 0.0047	-0.0173	0.0011	-0.33
8.7	340.24 ± 4.97	0.0154 ± 0.0080	-0.0081 ± 0.0047	-0.0173	0.0011	-0.32
8.8	340.24 ± 4.97	0.0154 ± 0.0080	-0.0081 ± 0.0047	-0.0173	0.0011	-0.31
8.9	340.24 ± 4.97	0.0153 ± 0.0080	-0.0081 ± 0.0047	-0.0173	0.0011	-0.34
9.0	340.23 ± 4.97	0.0151 ± 0.0080	-0.0079 ± 0.0047	-0.0171	0.0013	-0.46
9.1	340.23 ± 4.97	0.0151 ± 0.0080	-0.0079 ± 0.0047	-0.0171	0.0013	-0.47
9.2	340.23 ± 4.97	0.0153 ± 0.0080	-0.0080 ± 0.0047	-0.0172	0.0012	-0.36
9.3	340.22 ± 4.97	0.0169 ± 0.0080	-0.0086 ± 0.0047	-0.0178	0.0005	0.48
9.4	340.22 ± 4.97	0.0170 ± 0.0080	-0.0087 ± 0.0047	-0.0178	0.0005	0.52
9.5	340.22 ± 4.97	0.0171 ± 0.0080	-0.0088 ± 0.0047	-0.0179	0.0004	0.59
9.6	340.22 ± 4.97	0.0173 ± 0.0080	-0.0088 ± 0.0047	-0.0180	0.0003	0.65
9.7	340.24 ± 4.97	0.0173 ± 0.0080	-0.0090 ± 0.0047	-0.0182	0.0001	0.71
9.8	340.24 ± 4.97	0.0172 ± 0.0080	-0.0091 ± 0.0047	-0.0182	0.0001	0.70
9.9	340.24 ± 4.97	0.0172 ± 0.0080	-0.0091 ± 0.0047	-0.0182	0.0001	0.68
10.0	340.24 ± 4.97	0.0170 ± 0.0080	-0.0090 ± 0.0047	-0.0181	0.0001	0.60

Table 20: LMM summary results for values values of δ=[5.1, 10.0] - sequences produced after shuffling the order of whistles' duration

				Confidence intervals		
δ	Intercept	Position ± SE	Sequence Length ± SE	Lower	Upper	AICd
0.1	341.57 ± 5.10	-3.6557 ± 1.3624	-0.4543 ± 0.9651	-2.3458	1.4372	14.2419
0.2	342.11 ± 5.07	-1.4765 ± 1.0873	-0.9269 ± 0.7682	-2.4326	0.5788	7.1566
0.3	341.63 ± 5.06	-0.6656 ± 0.9557	-0.6930 ± 0.6738	-2.0136	0.6276	1.0978
0.4	342.11 ± 5.05	0.2515 ± 0.8599	-1.0862 ± 0.6066	-2.2751	0.1027	1.0863
0.5	341.22 ± 5.05	0.4999 ± 0.7903	-0.6385 ± 0.5586	-1.7334	0.4564	-2.6346
0.6	341.01 ± 5.04	0.8360 ± 0.7303	-0.6019 ± 0.5156	-1.6124	0.4087	-2.4320
0.7	340.73 ± 5.04	1.0385 ± 0.6877	-0.5232 ± 0.4861	-1.4759	0.4295	-1.7200
0.8	340.48 ± 5.04	1.2913 ± 0.6505	-0.4826 ± 0.4592	-1.3827	0.4175	0.1817
0.9	340.40 ± 5.04	1.6161 ± 0.6180	-0.5494 ± 0.4362	-1.4044	0.3056	3.5193
1.0	340.22 ± 5.04	1.7113 ± 0.5910	-0.5150 ± 0.4170	-1.3322	0.3023	5.6869
1.1	340.24 ± 5.03	1.7211 ± 0.5687	-0.5394 ± 0.4009	-1.3251	0.2463	6.4112
1.2	339.90 ± 5.03	1.4469 ± 0.5489	-0.3324 ± 0.3867	-1.0902	0.4255	4.9625
1.3	339.99 ± 5.03	1.5559 ± 0.5313	-0.4106 ± 0.3744	-1.1445	0.3232	6.4683
1.4	340.19 ± 5.03	1.4494 ± 0.5152	-0.4577 ± 0.3627	-1.1687	0.2532	4.9769
1.5	340.34 ± 5.03	1.6113 ± 0.5018	-0.5710 ± 0.3538	-1.2644	0.1223	7.1641
1.6	340.16 ± 5.03	1.6518 ± 0.4892	-0.5296 ± 0.3449	-1.2057	0.1465	8.8427
1.7	340.07 ± 5.03	1.7373 ± 0.4768	-0.5367 ± 0.3362	-1.1957	0.1223	11.1821
1.8	340.15 ± 5.03	1.6809 ± 0.4661	-0.5487 ± 0.3289	-1.1933	0.0958	10.5378
1.9	340.20 ± 5.03	1.6235 ± 0.4557	-0.5517 ± 0.3215	-1.1818	0.0784	9.9665
2.0	340.20 ± 5.03	1.5715 ± 0.4461	-0.5367 ± 0.3149	-1.1539	0.0805	9.6236
2.1	340.13 ± 5.03	1.6197 ± 0.4361	-0.5372 ± 0.3076	-1.1401	0.0658	11.3297
2.2	340.06 ± 5.03	1.8020 ± 0.4269	-0.5858 ± 0.3013	-1.1762	0.0047	15.9522
2.3	339.94 ± 5.03	1.8800 ± 0.4195	-0.5812 ± 0.2961	-1.1616	-0.0008	18.9540
2.4	340.18 ± 5.03	2.0333 ± 0.4106	-0.7137 ± 0.2894	-1.2810	-0.1465	22.6613
2.5	340.02 ± 5.03	2.0521 ± 0.4032	-0.6766 ± 0.2841	-1.2334	-0.1198	24.8670
2.6	340.10 ± 5.03	1.9515 ± 0.3955	-0.6689 ± 0.2785	-1.2147	-0.1231	22.7297
2.7	339.96 ± 5.03	1.9406 ± 0.3879	-0.6324 ± 0.2728	-1.1669	-0.0978	24.0514
2.8	339.78 ± 5.03	1.8368 ± 0.3821	-0.5484 ± 0.2686	-1.0749	-0.0218	22.8601
2.9	339.88 ± 5.03	1.8559 ± 0.3761	-0.5900 ± 0.2645	-1.1084	-0.0715	23.5643
3.0	339.57 ± 5.03	1.8146 ± 0.3704	-0.4964 ± 0.2604	-1.0069	0.0140	24.9298
3.1	339.49 ± 5.03	1.7040 ± 0.3644	-0.4400 ± 0.2561	-0.9420	0.0620	23.0052
3.2	339.63 ± 5.03	1.5594 ± 0.3592	-0.4283 ± 0.2526	-0.9234	0.0667	18.6835
3.3	339.69 ± 5.03	1.4535 ± 0.3542	-0.4093 ± 0.2489	-0.8971	0.0785	16.0694
3.4	339.58 ± 5.03	1.4311 ± 0.3498	-0.3765 ± 0.2459	-0.8585	0.1055	16.5121
3.5	339.42 ± 5.03	1.4666 ± 0.3458	-0.3525 ± 0.2435	-0.8297	0.1247	18.8305
3.6	339.45 ± 5.03	1.4040 ± 0.3408	-0.3411 ± 0.2397	-0.8109	0.1288	17.4747
3.7	339.35 ± 5.03	1.4013 ± 0.3359	-0.3212 ± 0.2363	-0.7844	0.1419	18.5210
3.8	339.40 ± 5.03	1.3868 ± 0.3312	-0.3335 ± 0.2328	-0.7898	0.1228	18.2818
3.9	339.40 ± 5.03	1.3481 ± 0.3268	-0.3214 ± 0.2298	-0.7718	0.1289	17.6926
4.0	339.36 ± 5.03	1.3187 ± 0.3227	-0.3046 ± 0.2269	-0.7493	0.1400	17.5575
4.1	339.56 ± 5.02	1.2399 ± 0.3189	-0.3247 ± 0.2242	-0.7643	0.1148	14.5580
4.2	339.39 ± 5.02	1.2530 ± 0.3150	-0.2956 ± 0.2215	-0.7297	0.1384	16.2579
4.3	339.07 ± 5.03	1.1904 ± 0.3117	-0.2018 ± 0.2193	-0.6317	0.2280	16.9862
4.4	339.09 ± 5.03	1.2804 ± 0.3079	-0.2453 ± 0.2166	-0.6698	0.1792	19.9062
4.5	339.20 ± 5.03	1.3701 ± 0.3045	-0.3085 ± 0.2142	-0.7284	0.1113	22.3777
4.6	339.22 ± 5.03	1.3687 ± 0.3011	-0.3165 ± 0.2118	-0.7317	0.0987	22.6535
4.7	339.34 ± 5.03	1.3611 ± 0.2981	-0.3422 ± 0.2098	-0.7534	0.0690	22.0004
4.8	339.33 ± 5.03	1.2985 ± 0.2945	-0.3214 ± 0.2072	-0.7275	0.0847	20.4096
4.9	339.41 ± 5.03	1.2932 ± 0.2918	-0.3392 ± 0.2054	-0.7418	0.0635	20.0812
5.0	339.44 ± 5.03	1.2982 ± 0.2891	-0.3505 ± 0.2037	-0.7497	0.0488	20.4104

Table 21: LMM summary results for values of δ=[0.1, 5.0] - sequences produced after shuffling the order of silences between the whistles and the order the whistles' duration

				Confidence intervals		
δ	Intercept	Position ± SE	Sequence Length ± SE	Lower	Upper	AICd
5.1	339.47 ± 5.03	1.2803 ± 0.2860	-0.3518 ± 0.2014	-0.7466	0.0430	20.0883
5.2	339.36 ± 5.03	1.2705 ± 0.2830	-0.3296 ± 0.1993	-0.7203	0.0610	20.8120
5.3	339.30 ± 5.02	1.1185 ± 0.2794	-0.2602 ± 0.1966	-0.6456	0.1251	16.6171
5.4	339.35 ± 5.02	1.0653 ± 0.2765	-0.2522 ± 0.1946	-0.6336	0.1292	14.9684
5.5	339.37 ± 5.02	1.0631 ± 0.2742	-0.2572 ± 0.1931	-0.6357	0.1213	15.0327
5.6	339.36 ± 5.02	1.1056 ± 0.2717	-0.2749 ± 0.1914	-0.6501	0.1002	16.7335
5.7	339.16 ± 5.02	1.0566 ± 0.2692	-0.2181 ± 0.1897	-0.5899	0.1537	16.7091
5.8	339.16 ± 5.02	1.0734 ± 0.2664	-0.2267 ± 0.1875	-0.5943	0.1409	17.6663
5.9	339.13 ± 5.02	1.0637 ± 0.2639	-0.2209 ± 0.1859	-0.5853	0.1435	17.8029
6.0	339.15 ± 5.02	1.0461 ± 0.2618	-0.2208 ± 0.1846	-0.5825	0.1410	17.2863
6.1	339.29 ± 5.02	0.9970 ± 0.2594	-0.2299 ± 0.1830	-0.5885	0.1288	15.0349
6.2	339.31 ± 5.02	0.9476 ± 0.2567	-0.2175 ± 0.1809	-0.5722	0.1371	13.5995
6.3	339.36 ± 5.02	0.9508 ± 0.2546	-0.2301 ± 0.1795	-0.5818	0.1216	13.6476
6.4	339.33 ± 5.02	0.9648 ± 0.2521	-0.2319 ± 0.1777	-0.5801	0.1163	14.5827
6.5	339.36 ± 5.02	0.9516 ± 0.2497	-0.2345 ± 0.1759	-0.5792	0.1102	14.2569
6.6	339.34 ± 5.02	0.9514 ± 0.2473	-0.2317 ± 0.1740	-0.5728	0.1094	14.7133
6.7	339.38 ± 5.02	0.9039 ± 0.2453	-0.2221 ± 0.1727	-0.5605	0.1163	13.1037
6.8	339.37 ± 5.02	0.9350 ± 0.2432	-0.2342 ± 0.1712	-0.5697	0.1014	14.4705
6.9	339.21 ± 5.02	0.8988 ± 0.2413	-0.1936 ± 0.1699	-0.5266	0.1394	14.3806
7.0	339.14 ± 5.02	0.9138 ± 0.2395	-0.1892 ± 0.1687	-0.5198	0.1415	15.5620
7.1	339.15 ± 5.02	0.9341 ± 0.2377	-0.2013 ± 0.1676	-0.5297	0.1272	16.4328
7.2	339.10 ± 5.02	0.9144 ± 0.2361	-0.1860 ± 0.1665	-0.5124	0.1403	16.2576
7.3	339.17 ± 5.02	0.9107 ± 0.2345	-0.1981 ± 0.1654	-0.5223	0.1262	15.8727
7.4	339.17 ± 5.02	0.9667 ± 0.2325	-0.2242 ± 0.1640	-0.5457	0.0972	18.2332
7.5	339.27 ± 5.02	0.9095 ± 0.2302	-0.2194 ± 0.1623	-0.5375	0.0988	15.7835
7.6	339.20 ± 5.02	0.8696 ± 0.2287	-0.1924 ± 0.1614	-0.5087	0.1239	14.9275
7.7	339.20 ± 5.02	0.8327 ± 0.2270	-0.1798 ± 0.1603	-0.4939	0.1344	13.7705
7.8	339.16 ± 5.02	0.8902 ± 0.2255	-0.1985 ± 0.1593	-0.5108	0.1137	16.3179
7.9	339.18 ± 5.02	0.8425 ± 0.2237	-0.1828 ± 0.1580	-0.4923	0.1268	14.6965
8.0	339.08 ± 5.02	0.8406 ± 0.2223	-0.1672 ± 0.1570	-0.4750	0.1406	15.4419
8.1	339.20 ± 5.02	0.9036 ± 0.2203	-0.2150 ± 0.1555	-0.5198	0.0898	17.4189
8.2	339.23 ± 5.02	0.9015 ± 0.2189	-0.2202 ± 0.1546	-0.5231	0.0827	17.3706
8.3	339.27 ± 5.02	0.8981 ± 0.2173	-0.2279 ± 0.1534	-0.5287	0.0729	17.1957
8.4	339.28 ± 5.02	0.8320 ± 0.2158	-0.2033 ± 0.1525	-0.5021	0.0955	14.7146
8.5	339.23 ± 5.02	0.8150 ± 0.2146	-0.1894 ± 0.1518	-0.4868	0.1081	14.5028
8.6	339.14 ± 5.02	0.7930 ± 0.2132	-0.1662 ± 0.1508	-0.4618	0.1294	14.4329
8.7	339.07 ± 5.02	0.8379 ± 0.2121	-0.1758 ± 0.1501	-0.4701	0.1184	16.7776
8.8	339.15 ± 5.02	0.8366 ± 0.2101	-0.1898 ± 0.1486	-0.4811	0.1016	16.5267
8.9	339.16 ± 5.02	0.8301 ± 0.2084	-0.1907 ± 0.1474	-0.4796	0.0983	16.4402
9.0	339.22 ± 5.02	0.8164 ± 0.2062	-0.1957 ± 0.1456	-0.4811	0.0898	15.8875
9.1	339.22 ± 5.02	0.7744 ± 0.2046	-0.1796 ± 0.1445	-0.4628	0.1036	14.4199
9.2	339.17 ± 5.02	0.7801 ± 0.2036	-0.1753 ± 0.1439	-0.4573	0.1068	15.0844
9.3	339.27 ± 5.02	0.7576 ± 0.2023	-0.1819 ± 0.1431	-0.4623	0.0985	13.7675
9.4	339.29 ± 5.02	0.7583 ± 0.2008	-0.1881 ± 0.1420	-0.4664	0.0901	13.8451
9.5	339.28 ± 5.02	0.7236 ± 0.1993	-0.1721 ± 0.1410	-0.4484	0.1041	12.7565
9.6	339.35 ± 5.02	0.6993 ± 0.1976	-0.1731 ± 0.1397	-0.4468	0.1007	11.6820
9.7	339.28 ± 5.02	0.6878 ± 0.1961	-0.1597 ± 0.1385	-0.4313	0.1118	11.7984
9.8	339.34 ± 5.02	0.6986 ± 0.1949	-0.1740 ± 0.1377	-0.4440	0.0959	12.0480
9.9	339.30 ± 5.02	0.6632 ± 0.1940	-0.1537 ± 0.1371	-0.4225	0.1150	11.0165
10.0	339.31 ± 5.03	0.6401 ± 0.1927	-0.1452 ± 0.1362	-0.4121	0.1217	10.3060

Table 22: LMM summary results for values values of δ=[5.1, 10.0] - sequences produced after shuffling the order of silences between the whistles and the order the whistles' duration

D | Shuffling algorithm

Here we show how the algorithm works in detail. As mentioned in the Methods section of chapter 3, the input of the algorithm is the list of whistle durations and the list of silent gaps between these whistles. For facilitating the understanding of each shuffling output, we use as an example words in the place of whistle durations. This way, it is easier and faster to comprehend the effect of the shufflings in the produced sequences.

Assume the following list of words W, which have occurred in that order with the silent gaps between them: :

$$W = [\text{'The', 'mind', 'was', 'dreaming', 'The', 'world', 'was', 'its', 'dream'}]$$

$$S = [0.2, 0.1, 0.12, 1.3, 0.1, 0.4, 0.1, 0.1]$$

Assume that we consider that the elements of list W belong to the same sequence, if the silent gap between them (δ) is less than 0.5 seconds. Applying the steps of algorithm 1, the output sequences are the following:

$$F = ([\text{'The', 'mind', 'was', 'dreaming'}], [\text{'The', 'world', 'was', 'its', 'dream'}])$$

Shuffling silences S

For the first shuffling scenario, we maintain the list W as it is, and we randomize the order of the silent gaps in S and we transform it into S_s.

$$S_s = [0.1, 1.3, 0.1, 0.4, 0.2, 0.1, 0.1, 0.12]$$

Applying the steps of algorithm 1, but this time with input S_1, W and δ, the output sequences are different:

$$F_s = ([\text{'The', 'mind'}], [\text{'was', 'dreaming', 'The', 'world', 'was', 'its', 'dream'}])$$

In this case of shuffling, the obtained sequences differ compared to the ones in F, as the silences order indicated that the break point should be different.

Shuffling durations W

The case of shuffled durations (words for this particular example), maintains the order of the silences list S and shuffles the order of W. A shuffled version of W is the following:

$$W_d = [\text{'The', 'was', 'dream', 'dreaming', 'its', 'The', 'mind', 'was', 'world'}]$$

We apply again the same steps of algorithm 1, using as input W_d, S and δ, and we obtain:

$$F_d = ([\text{'The', 'was', 'dream', 'dreaming'}], [\text{'its', 'The', 'mind', 'was', 'dream'}])$$

Again, the sequences we obtain are different from the initial ones. However, notice that in this shuffling case, the sequences will contain the same number of elements, as the thresholds that indicate where a sequence should stop have the same order as in the initial scenario. The difference is the contents of each sequence.

Shuffling both silences and durations

For the final case, we are shuffling both S and W, and we use their shuffled versions as input to obtain the final set of sequences. Below, S_b and W_b are the shuffled inputs and F_b contains the final sequences. The obtained results give sequences that have changed entirely in order and in size.

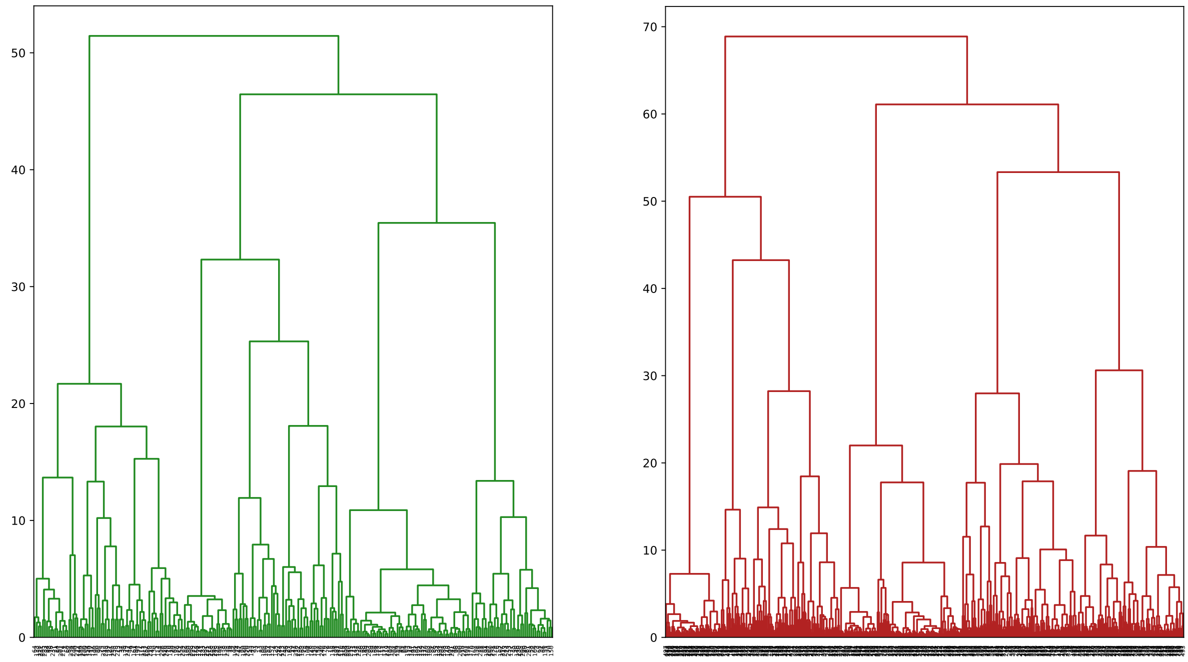
$$W_b = [\text{'was', 'dreaming', 'The', 'mind', 'was', 'The', 'world', 'its', 'dream'}]$$

$$S_b = [0.1, 1.3, 0.4, 0.12, 0.1, 0.1, 0.1, 0.2]$$

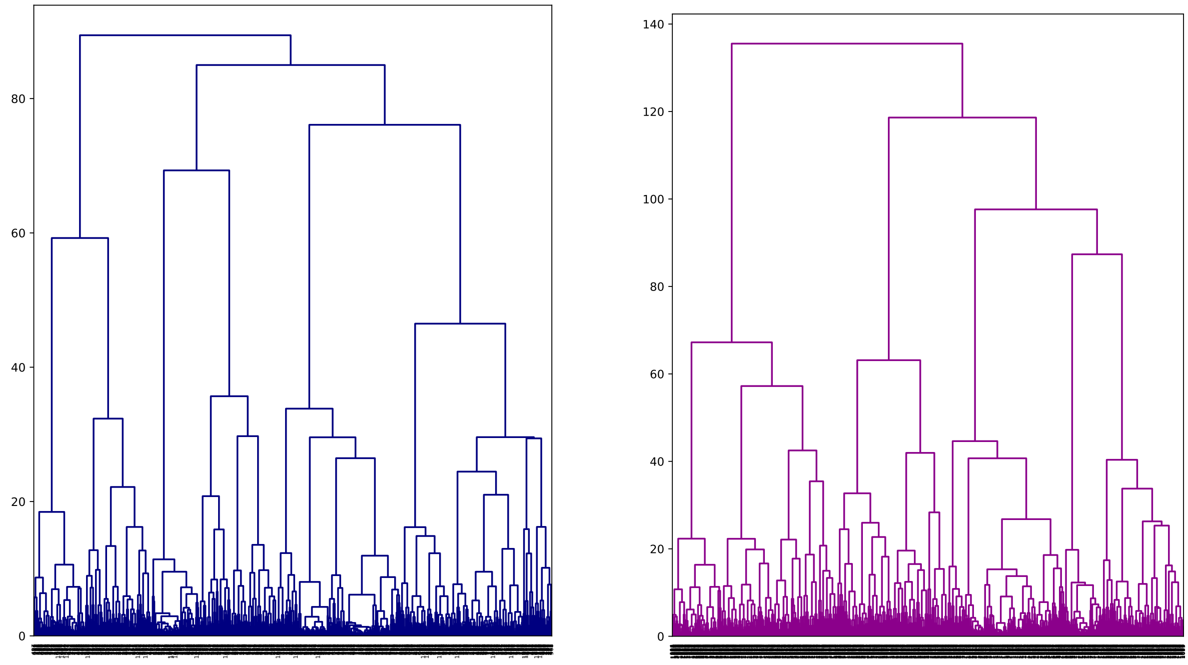
$$F_b = ([\text{'was', 'mind'}], [\text{'The', 'mind', 'was', 'The', 'world', 'its', 'dream'}])$$

E | Hierarchical clustering

Visualizations of dendrograms obtained after applying the two-phase clustering of chapter 4.3.



(a) $k_1 = 256$ (b) $k_1 = 512$



(c) $k_1 = 1024$ (d) $k_1 = 2048$

Figure 17: Dendrogram visualizations of hierarchical linkage. Each plot depicts the resulted dendrogram where we used as input the prototypes from the phase one clustering phase, using the corresponding k_1 parameter.