Author for correspondence:
Catherine Hobaiter
e-mail: clh42@st-andrews.ac.uk

THE ROYAL SOCIETY PUBLISHING

Variable expression of linguistic laws in ape gesture: a case study from chimpanzee sexual solicitation

Alexandra Safryghin[1], Catharine Cross[1], Brittany Fallon[1], Raphaela Heesen[3], Ramon Ferrer-i-Cancho[2,†] and Catherine Hobaiter[1,4,†]

[1]School of Psychology and Neuroscience, University of St Andrews, St Andrews, Fife, UK
[2]Complexity and Quantitative Linguistics Laboratory, Laboratory for Relational Algorithmics, Complexity, and Learning Research Group, Departament de Ciències de la Computació, Universitat Politècnica de Catalunya, 08034 Barcelona, Catalonia, Spain
[3]Department of Psychology, Durham University, Durham, UK
[4]Budongo Conservation Field Station, Masindi, Uganda

CC, 0000-0001-8110-8408; RH, 0000-0002-8730-1660; CH, 0000-0002-3893-0524

Two language laws have been identified as consistent patterns shaping animal behaviour, both acting on the organizational level of communicative systems. Zipf's law of brevity describes a negative relationship between behavioural length and frequency. Menzerath's law defines a negative correlation between the number of behaviours in a sequence and average length of the behaviour composing it. Both laws have been linked with the information-theoretic principle of compression, which tends to minimize code length. We investigated their presence in a case study of male chimpanzee sexual solicitation gesture. We failed to find evidence supporting Zipf's law of brevity, but solicitation gestures followed Menzerath's law: longer sequences had shorter average gesture duration. Our results extend previous findings suggesting gesturing may be limited by individual energetic constraints. However, such patterns may only emerge in sufficiently large datasets. Chimpanzee gestural repertoires do not appear to manifest a consistent principle of compression previously described in many other close-range systems of communication. Importantly, the same signallers and signals were previously shown to adhere to these laws in subsets of the repertoire when used in play; highlighting

1. Introduction

Over the past 100 years, important statistical regularities have been described across human languages and in other communicative systems such as genomes, proteins and animal vocal and gestural communication [1–11]. These regularities are hypothesized to be manifestations of the information-theoretic principle of compression [9,12]. Compression is a particular case of the principle of least effort [13]—a principle that promotes the outcome that requires the least amount of energy to produce or achieve—and thereby promotes coding efficiency [14]. In communication, compression is expressed as pressure towards reducing the energy needed to compose a code but limited by the need to retain the critical information in the transmission [15,16].

Among the statistical patterns predicted by compression at different levels of organization, Zipf's law of brevity and Menzerath's law have been at the centre of recent attention in studies of human and non-human communication. Zipf's law of brevity is the tendency for more frequent words to be shorter in length [13,17], and is generalized as the tendency for more frequent elements of many kinds (e.g. syllables, words, calls) to be shorter or smaller [14]—with similar patterns found at different levels of analysis, for example in speech at the level of words [17] syllables [18], and phonemes [4]. As well as being found in human spoken, signed and written languages [2–4,8,10], Zipf's law of brevity has been identified in the short-range communication of diverse taxa: dolphins [16], bats [19], penguins [20], hyraxes [21] and various primates (macaques: [22]; marmosets: [23]; gibbons: [24]; *Indri indri*: [25]), as well as in genomes [7].

At the level of constructs, Menzerath's law states that '*the greater the whole, the smaller its constituents*' [6,26,27]; for example longer sentences have words of shorter average length, and words with more syllables contain syllables of shorter length. Menzerath's law (and its mathematical expression known as the Menzerath-Altman's law) has been identified in human spoken and signed languages [26,28], genomes [29,30], music [31], and in the communication of dolphins [16], penguins [20] and primates (geladas: [32]; chimpanzees: [33,34]; gibbons: [24,35]; gorillas: [36]; *Indri indri*: [25]). While many studies focused on vocal communication, several have now explored these statistical regularities in gestural and signed domains. For example, the use of Swedish Sign Language in (semi-)spontaneous conversation was found to follow a pattern of more frequently used signs being shorter in duration [3]. Zipf's law of brevity was also found in fingerspelling, with a negative relationship between mean fingerspelled sign duration and frequency [3]. Similarly, Czech sign language was found to follow Menzerath's law [28]. Work in non-human gesture has, to date, been more focused on context-specific signal usage, for example Zipf's law of brevity was found in the surface behaviour of dolphins (such as tail-slapping; [37]) but not in the overall repertoire of play gestures of chimpanzees, where it was only present in subsets, although these gestures did follow Menzerath's law [34].

Chimpanzee gestural communication represents a powerful non-human model in which to explore compression and language laws. Apes have large repertoires of over 70 distinct gesture types [38]; as compared to vocal communication, gestural repertoires are larger and are more flexibly deployed, with individual gesture types used to achieve multiple goals [39–42]. Gestures are also used intentionally, i.e. to reach social goals by influencing the receivers' behaviour or understanding [41,43–45]), and flexibly across contexts [40–42]. Nevertheless, Heesen *et al.*'s [34] results support an increasingly diverse range of findings that show variation in the extent and expression of language laws, suggesting that while they appear statistically universal there is room for exceptions and/or variation in patterning at different levels of the communicative construct [9].

Although a lack of evidence supporting Zipf's law of brevity has been previously reported (e.g. European heraldry: [46]; computer-based neural networks: [47]), these remain rare exceptions, and in non-human animal communication have typically only been reported in long-distance vocal communication (e.g. gibbon song: [35]; bats: [19]; although cf. female hyrax calls: [21]) where the impact of distance on signal transmission fidelity may have a particularly strong effect on the costs of compression [9,14,32]. Thus, at present, the repertoire-level absence of Zipf's law of brevity in chimpanzee gesture remains a conundrum.

One explanation for a repertoire-level absence of Zipf's law of brevity—as seen in some long-distance signals—is that the context in which signals are produced may impact the emergence and expression of

these patterns. Specifically, in the case of chimpanzee gestures, the absence of a pattern resembling Zipf's law of brevity may result from the use of gestures produced during play. Expressions of linguistic laws in biological systems reflect pressures that shape efficient energy expenditure [9]. Play is produced when there is an excess of time and energy [48–50], thus, the energetic need to reduce signal effort through increased compression may be limited. As a result, it remains unclear whether the failure of Zipf's law of brevity in chimpanzee gesture was owing to the use of gestures from within play, or whether it reflects a system-wide characteristic.

In both signed languages and human gesturing, distinctions are made between different components of their production. First there is the *preparation* of the signal, then the *action stroke* which represents the movement that defines the gesture as of a particular type; an individual can then choose to further *hold* the stroke or repeat it, until they decide to stop gesturing and return the limb to rest during *recovery* from the gestural action [51]. For example: in a reach gesture this would correspond to the movement of the hand into position (*preparation*), the extension of the arm and hand towards the recipient (*action stroke*), the (optional) maintenance of the extension (*hold*), and finally the return of the hand and arm to a resting state (*recovery*). All four of these phases require some energetic investment to produce, but there may be variation across them, and aspects such as preparation and recovery may be nearly, or entirely, absent where several gestures are strung together. In some gesture types, their production does not include a *hold* phase (e.g. hit, jump, throw object); we term these *fixed* duration gestures, as the duration of their expression is relatively constrained across instances of production. Other gesture types can include a *hold* phase (for example: reach, object shake, swing) which may or may not be present, and, where present, may vary substantially in length; we term these *loose* duration gestures. There may be differences in the emergence of Zipf's and Menzerath's laws regarding the different components of gesture production. Menzerath's law acts from a proximate perspective on the building of communicative sequences in a specific communicative instance: for example, gestures produced in longer sequences may be shortened by variation the duration of components such as the shortening of the *hold* phase in loose gesture types. By contrast, Zipf's law acts on gesture types across instances of use—and as such may be less sensitive to the immediate context of production.

Another possible explanation for the variation in the emergency of compression in ape gesture is that the ability to detect linguistics laws, particularly where they are only subtly expressed, appears to require powerful datasets. The exploration of statistical patterns in human languages often employs corpora containing millions of data points (e.g. [52]). By contrast, in ape gesture, as in many studies of non-human communication, datasets are substantially smaller (in the thousands). In chimpanzee play, the large repertoire expressed limits the frequency with which particular gesture types are represented.

We address this open question in a case study of chimpanzee gestural communication in sexual solicitation. While gesture is relatively under-studied in this area, sexual solicitations have been contrasted with early descriptions of gesture from studies of captive ape play, as an example of gesture in a relatively more evolutionarily or biologically 'relevant' context for communication (in terms of associated risks and/or impact on reproduction) ([53]; cf. [40]). Chimpanzees, particularly male chimpanzees, employ prolific use of individual gestures and gesture sequences in sexual solicitations. As solicitations are often vigorous, chimpanzees incorporate regular use of gesture types that include both visual and audible information [53,54]. While a range of gesture types are employed, these are typically a smaller sub-set of the available repertoire—cf. play where the majority of gesture types are deployed. Successful gestures can lead directly to sexual behaviour, such as inspection or copulation, as well as to a consortship, in which the female follows the male away from other individuals in the group so that he maintains exclusive sexual access [55]. Both direct solicitation and consortship and are key strategies for individual fitness [55,56], and as such behaviour associated with them is probably subject to strong selective pressures. The energetic costs of lactation mean that adult female chimpanzees typically concieve only once every 4–5 years [57,58]. So while there are typically 60–80 individuals in a group, the operational sex ratio of available females in oestrus may be very small, and males show substantial variation in reproductive success [55,59]. Although highly important, the performance of sexual solicitations may come with significant costs: besides the energetic expenditure in producing these signals, there is a risk of potentially aggressive competition both from other males in their own community [55,60] as well as potentially lethal attacks from males in neighbouring groups [61]. For example, during consortships individuals may travel to the boundaries of their home area, increasing the risk of encounters with neighbouring individuals. Thus, there are substantial advantages to avoiding potential eavesdroppers within, and particularly outside of, one's community [62]. Therefore, on one hand individuals benefit from producing conspicuous energetic signals to attract females, often having to insist to secure mating; on
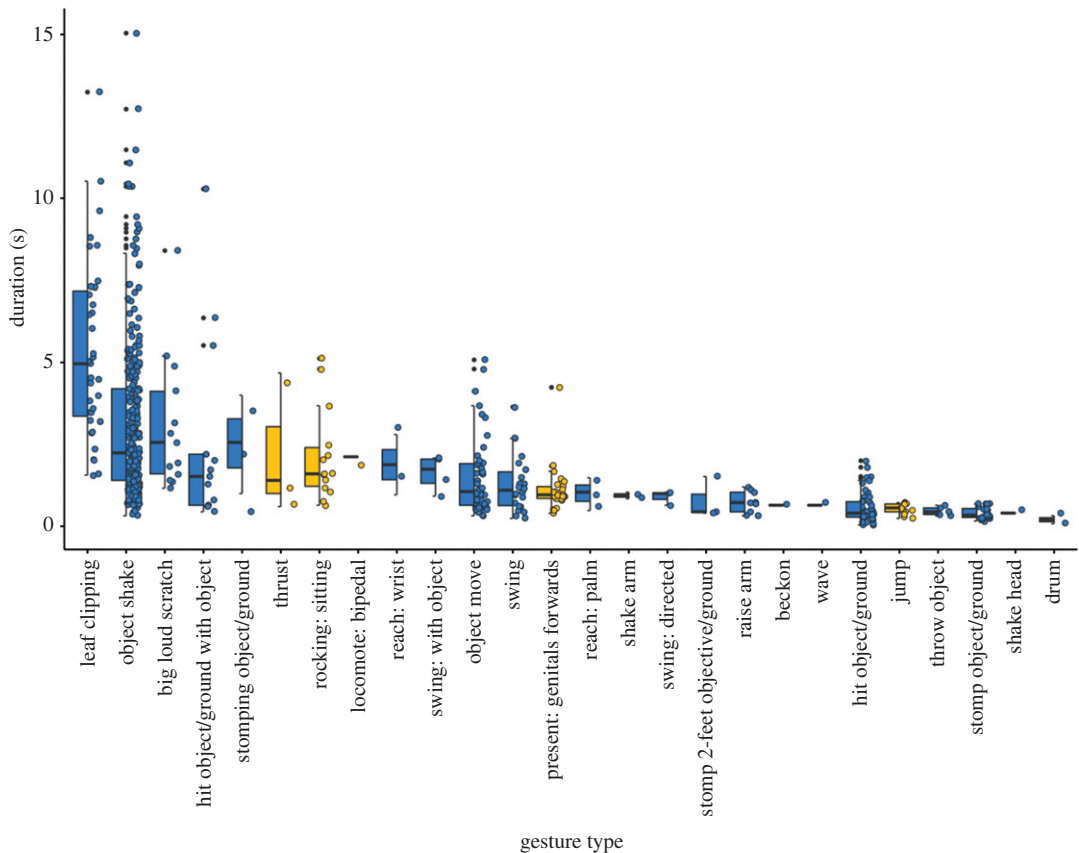
Figure 1. Distribution of gestural instances across the 26 gesture types detected with relative gesture duration. Boxplots represent the median (black bar), the interquartile range (boxes), maximum and minimum values excluding outliers (whiskers) and outliers (black dots). Points represent individual gestures. Whole-body gestures are indicated in yellow, manual gestures in blue.

the other, the production of highly conspicuous signals should be compressed to reduce the risks associated with competition from both within and outside the group.

To assess compression in the sexual solicitation gestures of wild male chimpanzees, we tested for patterns predicted by Zipf's law of brevity and Menzerath's law, both at the level of single gesture types and gesture sequences, respectively. To investigate Zipf's law of brevity and Menzerath's law we fitted two generalized linear mixed models. The first model explored the presence of Zipf's law assigning gesture duration as the response variable, proportion of gestures within the dataset and category of gesture (manual versus whole body) as fixed factors, and signaller's identity (ID), sequence ID and gesture type as random factors. The second model tested for Menzerath's law and had gesture duration as response variable, sequence size as a fixed factor, and proportion of whole-body gestures in the sequence (PWB), signaller ID, sequence ID and gesture type as random factors. We included information on the category of the gesture to allow for comparisons with human studies, in which gestures are mostly manual. We provide matched models that describe the patterns of expression both across (i) all males in our data, (ii) for a single prolific individual and (iii) for the remaining individuals. In doing so, we provide an initial assessment of the distribution of our findings across male chimpanzee gesturing in this context and provide an expanded assessment of compression in ape gestural communication.

2. Results

We measured $n = 560$ sexual solicitation gestures from 173 videos of 16 wild, habituated male East African chimpanzees (*Pan troglodytes schweinfurthii*) gesturing to 26 females. Within the 560 gestural instances (from now *tokens*), we identified 26 gesture types: 21 manual gestures and five whole-body gestures (figure 1; for definitions for full repertoire definitions see the electronic supplementary material, S1 and table S1) performed by 16 male chimpanzees aged 10–42 years old. On average, each

Table 1. Number of sequences composed of the same or different gesture types, listed according to sequence length. (Sequence length is defined by the number of gesture tokens present in the sequence.)

sequence length	the same type	different types	total number of sequences
1	NA	NA	244
2	24	58	82
3	1	20	21
4	0	3	3
5	1	6	7
6	0	2	2
total	26	89	

individual produced a median of 11.5 ± 70.7 gesture tokens (range 2–290). One male, Duane, was particularly prolific ($n = 290$ gesture tokens; other males 2–76). To provide context as to what extent our findings are generalizable, we provide matched analyses using both the full dataset and the dataset limited to Duane only. An analysis of the data excluding Duane is available in the electronic supplementary material.

Gesture token duration was measured via analysis of video data with a minimum unit of 0.04 s (one frame). Duration ranged from 0.04–15.04 s (median: 1.56 ± 2.35 s). If consecutive gesture tokens were performed with less than 1 s in between them, they were considered to form a sequence [34,63]. We detected a total of 377 sequences, with each male performing a median of 8 ± 44.54 sequences (range 1–181 sequences). Sequence length ranged from 1 to 6 tokens (table 1). For analyses of Menzerath's law we excluded 18 sequences for which we were unable to identify the duration of all the consecutive gesture tokens performed, resulting in the analysis of 359 sequences, containing a total of 530 gesture tokens. There were 244 sequences composed of a single token, the remaining 115 sequences had length $n > 1$. Of the 115 sequences analysed that were composed of two or more gesture tokens; 26 (23%) were formed by the repetition of the same gesture type, whereas the remaining 89 (77%) included more than one gesture type (table 1).

2.1. Do chimpanzee sexual solicitation gestures follow Zipf's law of brevity?

To test for Zipf's law we ran a Bayesian generalized linear model (Zipf-model), with the log of gesture duration as the response variable and the proportion of gesture type within the dataset as a fixed factor (see the electronic supplementary material, S2 for further detail). The gesture duration data was log-transformed following an analysis of data distribution. We included category of gesture as a control, and signaller ID, sequence ID and gesture type as random factors. The Zipf-model fitted the data better than a null model that did not include the proportion of gesture type as a fixed effect (leave-one-out (LOO) difference and s.d. $= -0.7 \pm 0.3$). For Zipf-model effects bulk effective sample size (ESS) and tail ESS were > 100 and $\hat{R} = 1$. However, the proportion of gesture type did not have a substantial effect on the duration of gestures (electronic supplementary material, S3 and table S5; $b = 0.90$, s.d. $= 1.26$, 95% credible intervals (CI) $[-1.25, 3.81]$; figure 2a). When testing the subset of data containing only the gestures produced by Duane, the full model and null model testing for Zipf's law showed similar fit (LOO difference: -0.1 ± 0.7; electronic supplementary material, S3 and table S6; figure 2b). Similarly, in the same analysis on data from all individuals except Duane, the full model was no different from the null model (LOO difference: -0.5 ± 0.5; electronic supplementary material, S3 and table S7).

2.2. Do chimpanzee sexual solicitation gesture sequences follow Menzerath's law?

To test for Menzerath's law we ran a second Bayesian model (Menzerath-model) with the log of the gesture duration as response variable, the sequence size as fixed factor, the PWB as a control, and the signaller ID and sequence ID as random factors. The Menzerath-model fitted the data better than the null model (LOO difference: -7.7 ± 4.1). All predictors had bulk ESS and tail ESS > 100 as well as $\hat{R}$ values $= 1$. Sequence size had a substantial negative effect on gesture duration within sequence (electronic supplementary material, S3 and table S8; $b = -0.18$, s.d. $= 0.04$, 95% CI $[-0.26, -0.11]$; figure 3a). Similar results were found when
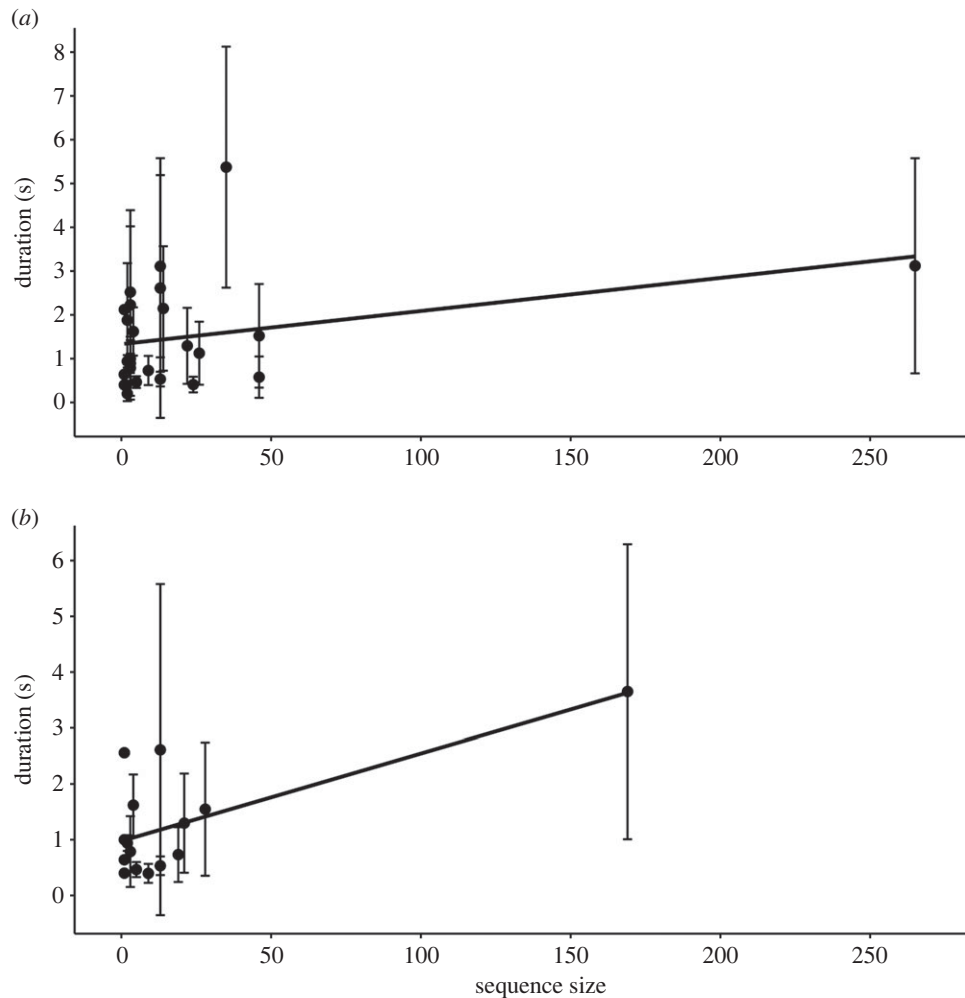
Figure 2. Relationship between frequency of occurrence and gesture duration for the full dataset (*a*) and Duane only data (*b*). Points represent the mean duration of each gesture type, with error bars showing the standard deviation from the mean. Black line indicates regression slope.

running the same Menzerath-model but limited to gestures produced by Duane: the full model fitted the data better than the null (LOO difference: -13.5 ± 4.5), all predictors had bulk ESS and tail ESS > 100, $\hat{R} = 1$ and sequence size had a substantial negative effect on gesture duration (electronic supplementary material, S3 and S9; figure 3*b*; $b = -0.23$, s.d. = 0.04, 95% CI [-0.31, -0.15]). By contrast, where Duane's data were excluded, the full model was similar to the null model, suggesting no clear pattern consistent with Menzerath's law (LOO difference: -0.3 ± 0.8; electronic supplementary material, S3 and table S10). Visual inspection of the data plotted per individual suggests that detection of a pattern consistent with Menzerath's law may be impacted by sample size (electronic supplementary material, S4).

We note that the sample size of sequences of four tokens or longer is smaller than those of one to three tokens (table 1), which may have contributed to the apparent tailing off of a clear relationship in figure 3*a*, *b*. In addition, longer sequences were formed of (i) a mix of *loose* and *fixed* duration gestures, or (ii) only *loose* duration gestures (see the electronic supplementary material, S5 and figures S5 and S6). Thus, the emergence of Menzerath's law could not be explained by a shift in preference from *fixed* to *loose* gestures with increasing sequence length.

3. Discussion

Chimpanzee sexual solicitation gestures did not follow Zipf's law of brevity: the frequency of gesture type within the dataset did not predict gesture duration in any of our samples. However, sequences of chimpanzee solicitation gestures did follow Menzerath's law: longer sequences of gestures were made up of gestures of shorter average length. Our dataset was limited both by its relatively small size (cf.
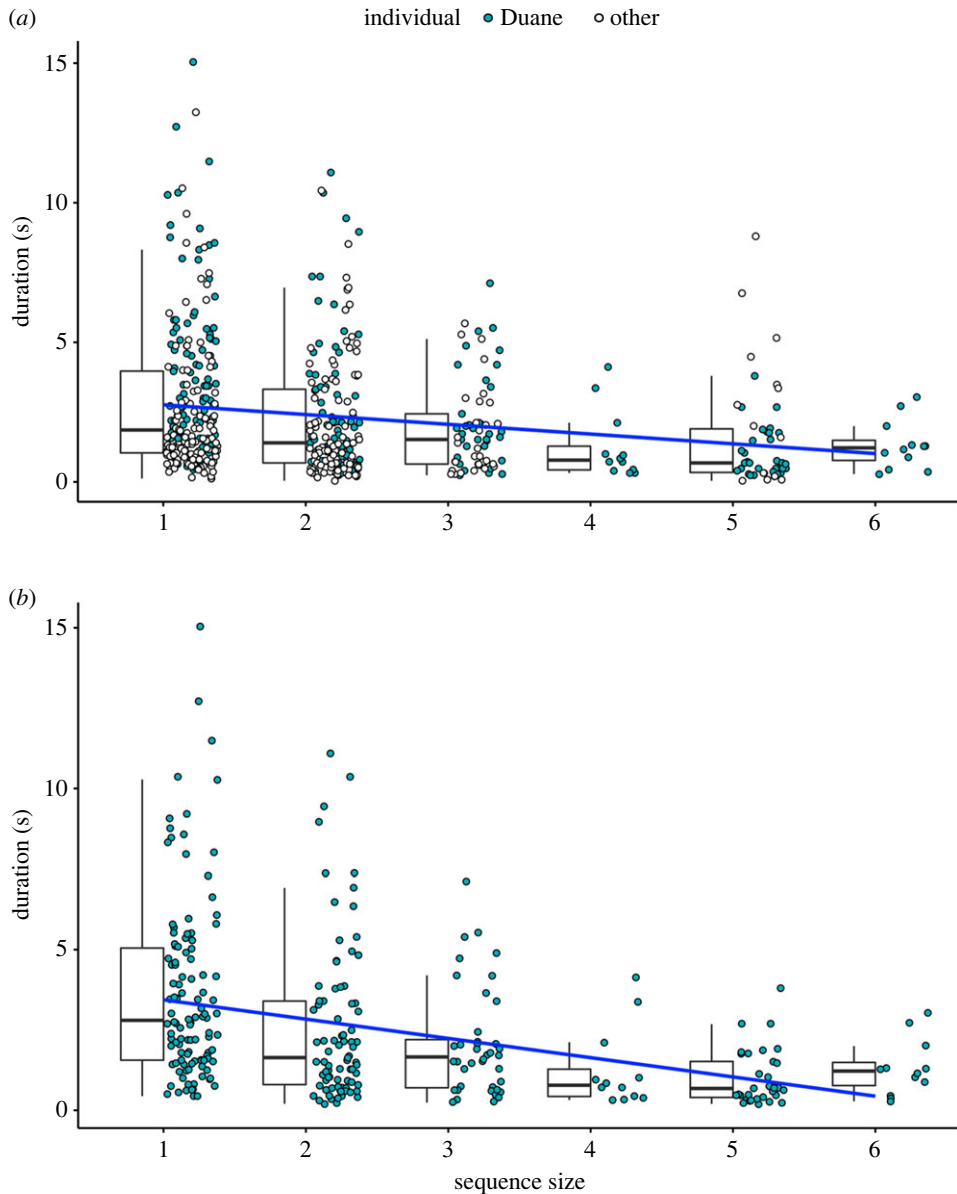
Figure 3. Relationship between sequence size and gesture duration for the full dataset (*a*) and Duane only data (*b*). Boxplots represent the median (black bar), the interquartile range (boxes), and maximum and minimum values excluding outliers (whiskers). Points represent individual gesture tokens, ordered by the length of the sequence they were performed in. Gestural tokens belonging to the individual Duane are indicated in light blue. White circles indicate gesture tokens belonging to all other individuals. Blue line indicates regression slope.

[34] on chimpanzee play gestures) and in its bias towards a single highly prolific individual (Duane). As a result, we consider it a case-study; however, the pattern was present in both Duane's data and in the full dataset, as well as in a range of alternative analyses (electronic supplementary material, S6). In the reduced dataset excluding Duane we did not find a pattern consistent with Menzerath's law; however, detection of the pattern may have been limited by the small sample size available in the remaining dataset.

These results represent a further absence of evidence in support of Zipf's law of brevity in great ape gestural communication [34] and support the wider finding that—unlike most other close-range systems of communication described to date—the expression of pressure for compression and efficiency may be variably expressed in ape gesture [3,9,14]. It particularly highlights that compression does not act on communicative systems uniformly: 20 of the 26 gesture types described here as used in sexual solicitations overlapped with those used in play [34]. Data were collected from the same community over the same period, and although both studies provided a null result when analysing the full

gestural repertoire, Zipf's law was found in subsets of the play gestures but not in the gestures when used in sexual solicitations. Moreover, when running traditional correlation analyses in which features such as signaller identity, or gesture type could not be controlled for, we found a tendency for an opposite Zipf's law pattern—particularly in manual gestures (electronic supplementary material, S6). Visual inspection of figure 3 shows the substantial variation in the duration of gestures across instances of communication, as well as an apparent decrease in a clear relationship between gesture duration and sequence size where sample size was small (such as for longer sequences). Together these findings suggest that the expression of these laws is nuanced by aspects of the communicative landscape in which they are deployed, and that large samples may be needed to detect sometimes subtle relationships. Future work could specifically explore variation in the detection of these patterns at different sample sizes, for example by randomized subsetting of sufficiently large datasets. As Semple *et al.* [9] suggest, apparent 'failures' may be of substantial assistance in exploring the boundaries of the theoretical framework of these laws, helping to define the characteristics that shape both their emergence and variation in their expression.

By contrast to vocal communication across primate species, in chimpanzee sexual solicitations 'inefficiency' in signalling effort by the signaller appears to be at times slightly favoured. However, these gestures appear to remain effective in terms of achieving the signaller's goal of successful communication in a context vital for reproductive success. Given the long inter-birth intervals and active mate guarding [64], chimpanzee paternity is often heavily biased towards higher-ranking individuals [59]. With so few opportunities to mate, sexual solicitations may represent one of the most evolutionarily important contexts in which chimpanzee gestures are produced. Where the costs of signal failure are high, there is a pressure against compression and towards redundancy, as in chimpanzees' use of gesture-vocal signal combinations in agonistic social interactions [62]. While there are examples of vocal communication systems used in biologically 'relevant' contexts that adhere to Zipf's brevity law [20], the benefits of successful communication to individual fitness in chimpanzee solicitation appear to outweigh the energetic costs associated with the production of a vigorous and conspicuous signal. Nevertheless, given that we see a relatively consistent expression of Menzerath's law across gesture use in sexual solicitation as in play, even the production of these prolonged and conspicuous signals appear to remain constrained by physiological mechanisms of gestural production. As for primate vocal communication [32,33], where breathing constraints and energetic demands of vocal production were considered drivers for the emergence of Menzerath's law patterns, increased muscular activity related to the production of sequences of gestures [65] could be a general limit on energetic investment. As a result, Menzerath's law appears to emerge across communicative contexts.

There are a number of potential reasons for why language laws appear variable in their expression within ape gesture. For example, we might be considering the wrong unit of analysis. In human speech, sign, and gesture—as in other communication systems—it is possible to consider the production of a 'unit' of communication at different levels. For example, while Zipf's law is clearly expressed in the duration of male rock hyrax vocalizations, it is not the case for female vocalizations where Zipf's law of brevity emerges when analysing call amplitude rather than duration [21]. Conversely, in Börstell *et al.* [3] research on Swedish Sign Language, Zipf's law of brevity seems to hold across sign categorization, fingerspelling and compounding. Interestingly, this study excluded the hold phase of a sign, limiting their analysis only to the more active stroke phase. The production of intentional gestures in apes are shaped not only by the signaller but by the interaction between signaller and recipient [38,66]. As a result, the duration of hold or repetition phase may be shaped by the immediate context of the specific interaction—for example, in waiting for a response by the recipient it may vary between being absent and very prolonged. By contrast, the *action stroke* of a sign or gesture is always present and represents the need to convey information in that gesture, i.e. to discriminate it from other gesture actions. In Swedish Sign Language a prolonged and repetitive feedback sign and prolonged turn-taking signs were the only two cases that diverged from the general Zipf's pattern, as they were both long in duration as well as being highly frequent [3]. Zipf's law acts on a signal 'type' in an individual's or species' repertoire—and it may be of interest to compare its expression across areas of gesture production that are more consistently produced across usage, such as the action stroke.

Research to date has typically focused on signal compression at the level of the communication system, but communication happens *in situ*. Signallers probably respond to pressures on signalling efficiency more broadly: an intense but time-limited investment in clear signalling may be more energetically efficient than the need to travel with a female for extended periods following a failed signal. A similar solicitation with a different audience may need to be produced rapidly and inconspicuously, as the detection of this activity by other males could be fatal [60]. In a recent human

study, pressures towards efficiency and accuracy were both required for Zipf's law of brevity to emerge in experimental communicative tasks between two participants [67]. Conversely, when participants were required to produce solely time-efficient versus solely accurate communicative signals no pattern emerged. The sexual solicitation context tested in our study may mirror the pattern seen in the time-efficient paradigm in the human study. In play, where urgency and time-efficiency may be less relevant, the same signals used by the same chimpanzees did show compression. While many vocalizations are relatively fixed [68,69], gestural flexibility (in goal and context—[39–42]) allows us to explore how compression acts within both specific instances of communication as well as on whole communication systems. To do so will require large longitudinal datasets in which it is possible to test both between-individual variation and within-individual variation across different gesture types and sequence lengths. Similarly, there remains substantial work needed to explore variation across different socio-ecological contexts of gesture use, for example in the social relationship between the signaller and recipient [66]. The use of redundancy within specific subsets of gestural repertoire, or within specific contexts of gesture demonstrates both the importance of compression in communicative systems in general, but also the flexibility present in each specific usage. In doing so, it highlights the importance of exploring the impact of individual and socio-ecological factors within wider patterns of compression in biological systems in evolutionary salient scenarios.

4. Methods

We measured $n = 560$ male-to-female sexual solicitation gestures from 173 videos recorded within a long-term study of chimpanzee gestural communication depicting 16 wild, habituated East African chimpanzees (*Pan troglodytes schweinfurthii*) from the Sonso community of the Budongo Forest Reserve in Uganda (1°35′ and 1°55′ N and 31°08′ and 31°42′ E), collected between December 2007 and February 2014. Observations were made between 7.30 and 16.300 with recording of gestures following a focal behaviour sampling approach [70]. Here, all social interactions were judged to have the potential for gesture, in practice any situation in which two chimpanzees were in proximity and not involved in solitary activities, were targeted. Where several potential opportunities to record co-occurred, preference was given to individuals from whom fewer data had been collected (with a running record of data collection maintained to facilitate these decisions).

During October 2007 to August 2009 a Sony Handycam (DCR-HC-55) was used. Here video was recorded on MiniDV tape. The challenges of filming wild chimpanzees in a visually dense rainforest environment meant that, at times, the start of gestural sequences was not captured on video. Where this occurred, it was dictated onto the end of the video and these sequences were not included in analysis. Similarly, sequences in which part of the sequence was obscured, for example where a chimpanzee moves through dense undergrowth, were also discarded. After 2009 video data were collected using Panasonic camcorders (V770, HC-VXF1) which have a 3-second pre-record feature that improves the ability to capture the onset of behaviour; however, the same procedure was used and any sequences where the onset of gesturing was not clearly captured continued to be discarded.

4.1. Sexual solicitation gestures

Sexual solicitation gestures were defined as those gestures given by a male towards a female with the goal of achieving sex, usually accompanied by the male having an erection and the female being in oestrus [41,53]. We included solicitations in the context of sexual consortship; here a male gestures in order to escort a female away from the group to maintain exclusive sexual access, which can occur prior to the peak of the female oestrus [55]. We restricted our analyses to male to female sexual solicitation, as female to male sexual solicitation attempts rarely involved sequences of gestures in this population. We further restricted analysis to solicitations by male individuals of at least 8 years old, as this is the minimum age of siring recorded in this community, limiting our signals to those on which there is more direct selective pressure.

4.2. Defining gesture types and tokens

In quantitative linguistics, word *types* are used to assess Zipf's law of brevity, whereas *tokens* are used to assess patterns conforming to Menzerath's law. To distinguish the two, consider the question:

Which witch was which?

The question is composed of 4 *tokens* (overall word count), and three different word *types*, (which, witch, was). Gesture *types* (see the electronic supplementary material, table S4 for a detailed repertoire description) were categorized according to the similarity of the gesture movement, which could be used either as a single instance or in a sequence; and each gestural instance represented an individual *token*.

Great apes deploy gestural sequences in two distinct forms [63]: one is the addition of further gestures following response waiting and is typically described as persistence (which may include elaboration). The second is the production of gestures in a 'rapid sequence'—here gestures are produced with less than 1 s between consecutive gesture tokens, and do not meet behavioural criteria for response-waiting occurring within a sequence (although it may occur at the end of it). As the expression of Menzerath's law is typically considered at the level of a unique sequence, rather than one generated through the addition of gestures in response to earlier failure, we limit our analyses here to rapid sequences only. Sequence length was quantified as the number of gesture tokens produced with less than 1 s between two consecutive gesture tokens; single gestures were coded as sequences of length one [34,63].

4.3. Gesture duration

Gesture duration was calculated using MPEG streamclip (v. 1.9.3beta). We measured gesture duration in frames, each lasting 0.04 s. Gestural 'units'—like many other signals—can be considered at different levels of analysis, for example: a word is composed of syllables, and syllables of phonemes. Gestures have been described as composed of preparation, action stroke, hold or repetition and recovery phase [51]. Here we follow previous work in [34] in defining the start of a gesture token as the initial movement of a part of the body required to produce the gesture. The end of a gesture token corresponded to (i) the cessation of the body movement related to gesture production, (ii) a change in body positioning if the gesture relied on body alignment, or (iii) the point at which the goal was fulfilled, and any further movement represented effective action (for example, locomotion or copulation). Where the expression of a gesture token did not include a full recovery (in which the body part involved is returned to a resting state), the end of a token was discriminated from subsequent tokens through (i) a change in gesture action, e.g. from a reach to a shake, (ii) a change in orientation or rhythm of a gesture action, hold or repetition, e.g. the rhythm or direction of an object shake is broken or changed [62].

4.4. Intra-observer reliability

Video-based coding offers the opportunity to conduct reliability measures. Intra-observer reliability was tested by randomizing the order of the videos and re-coding the duration of the gestures of every ninth clip, for a total of 75 gestures from 23 clips. We performed an intraclass correlation coefficient (ICC) test—class 3 with $n = 1$ rater [71]—which revealed very high agreement on gesture duration measurements (ICC = 0.995, $p < 0.001$). Unfortunately, an additional step of inter-observer reliability was not possible owing to the loss of the file that linked the original dataset to the videos from which data were extracted.

4.5. Statistical analysis

All data were analysed using R v. 4.0.0 and RStudio v. 1.2.5042 [72,73]. We fitted Bayesian generalized linear multivariate multilevel models using the 'brm' function from the 'brms' package [74] with minimally informative priors, 2000 iterations and three chains.

We ran a first model testing Zipf's law of brevity (Zipf-model), containing gesture token duration (s) as the response variable, the proportion of occurrences of a particular gesture type in the dataset (proportion) as a fixed effect, and gesture category (manual versus whole-body) as a control. We included signaller ID, sequence ID and gesture type as random effects. We include category as a variable here to allow for more direct comparison with previous work, which often excludes or differentiates non-manual signals, either in great ape gesture [34,75] or in signed languages and fingerspelling (e.g. [3]).

We tested Menzerath's law by running a second model (Menzerath-model) containing gesture token duration (s) as the response variable, sequence size (number of gesture tokens within the sequence) as a fixed factor, and the PWB as a control. We modelled signaller ID and sequence ID as random factors.

It was highlighted during the review process that the emergence of Menzerath's law may be an artifice created by the selection of *fixed*, as opposed to *loose*, duration gesture types when producing

longer sequences. To address this hypothesis, we produced histograms depicting the distribution of *loose* and *fixed* duration gestures within sequences at each sequence size. The majority of gesture types ($n = 20$ of total 26), and of gesture tokens ($n = 456$ of total 560) were of the *loose* gesture form, thus there were very few gesture sequences formed only of *fixed* gesture types. However, we further visually assessed the distributions of *fixed* gestures in sequences formed of only *fixed* gestures.

As our data may be particularly influenced by a single prolific individual (Duane) who contributed around half of the data, we assess the generalizability of our findings by replicating analyses conducted on the full dataset on a subset of the data containing only gestures by Duane as well as on a subset containing all but the prolific individual Duane. For the models testing Duane's data, signaller ID was removed from the random factors as it was no longer relevant (with the inclusion of only one individual). In order to avoid inflation of the dataset we include date as a random factor; which also allows us to avoid biasing the analysis towards particularly prolific days and control for within-individual consistency.

We ran full-null model comparisons using the LOO information criterion [76] 'loo_compare' function from the 'stan' package (v. 2.21.5; [77]) where Zipf's null model contained only the control variable 'category' and the random effects, whereas Menzerath's null model contained only the control variable PWB and the random effects. Prior to the Bayesian analysis we assessed data distribution using the 'fitdistr' package (v. 1.0–14; [78]). Following data inspection, we log-transformed gesture duration and average sequence duration as data from the response variable strongly skewed towards zero (for data inspection see the electronic supplementary material, S2).

Finally, previous work has frequently employed correlation and compression tests, which looks at whether the expected mean code length observed in the dataset is significantly smaller than a range of mean code lengths calculated via permutations, to test the mathematical theory behind both laws. In addition, we also fitted Bayesian generalized linear multivariate multilevel models with same number of iterations and chains as the previous models but having the median duration of each of the 26 gesture types as response variable, category of gesture as a fixed factor, as well as frequency of that gesture type as a predictor. These tests offer limited opportunities to control for potential confounds such as signaller identity and should be interpreted with caution in relatively small and variable datasets. We provide them in the electronic supplementary material, S6 to allow for comparison with previous work that analysed median durations with or without implementing generalized linear models (e.g. [4,36]).

References

1. Altmann G, Gerlach M. 2016 Statistical laws in linguistics. In *Creativity and universality in language* (eds MD Esposti, AF Pachet), pp. 7–26. Berlin, Germany: Springer International Publishing.

2. Bentz C, Ferrer-i-Cancho R. 2016 Zipf's law of abbreviation as a language universal. Leiden workshop on capturing phylogenetic algorithms for linguistics, Leiden, The Netherlands.

3. Börstell C, Hörberg T, Östling R. 2016 Distribution and duration of signs and parts of speech in Swedish Sign Language. *Sign Lang. Linguist.* **19**, 143–196. (doi:10.1075/sll.19.2.01bor)

4. Hernández-Fernández AG, Torre I, Garrido J-M, Lacasa L. 2019 Linguistic laws in speech: the case of Catalan and Spanish. *Entropy* **21**, 1153. (doi:10.3390/e21121153)

5. Köhler R, Altmann G, Piotrowski RG. (Eds.). 2005 *Quantitative linguistik/quantitative linguistics: Ein internationales handbuch/An international handbook*. Berlin, Germany: De Gruyter Mouton.

6. Menzerath P. 1954 *Die architektonik des deutschen wortschatzes*, vol. 3. Bohn, Germany: F. Dümmler.

7. Naranan S, Balasubrahmanyan VK. 2000 Information theory and algorithmic complexity: applications to linguistic discourses and DNA sequences as complex systems part I: efficiency of the genetic code of DNA. *J. Quant. Linguist.* **7**, 129–151. (doi:10.1076/0929-6174(200008)07:02;1-Z;FT129)

8. Sanada H. 2008 *Investigations in Japanese historical lexicology (revised edn)*. Göttingen, Germany: Peust & Gutschmidt.

9. Semple S, Ferrer-i-Cancho R, Gustison ML. 2022 Linguistic laws in biology. *Trends Ecol. Evol.* **37**, 53–66. (doi:10.1016/j.tree.2021.08.012)

10. Wang Y, Chen X. 2015 Structural complexity of simplified Chinese characters. In *Recent contributions to quantitative linguistics* (eds A Tuzzi, M Benesová, J Macutek), pp. 229–240. De Gruyter.

11. Zipf GK. 1936 *The psycho-biology of language: an introduction to dynamic philology*. Abingdon, UK: Routledge.

12. Ferrer-i-Cancho R, Bentz C, Seguin C. 2022 Optimal coding and the origins of Zipfian laws. *J. Quant. Linguist.* **29**, 165–194. (doi:10.1080/09296174.2020.1778387)

13. Zipf GK. 1949 *Human behavior and the principle of least effort: an introduction to human ecology*. Cambridge, MA: Addison-Wesley Press.

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

15. Cover TM, Thomas JA. 2006 *Elements of information theory*, 2nd edn. Hoboken, NJ: Wiley-Interscience.

16. Ferrer-i-Cancho R, Lusseau D, McCowan B. 2022 Parallels of human language in the behavior of bottlenose dolphins. *Linguist. Front.* **5**, 1–7. (doi:10.2478/lf-2022-0002)

17. Strauss U, Grzybek P, Altmann G. 2007 Word length and word frequency. In *Contributions to the science of text and language, vol. 31* (ed. P Grzybek), pp. 277–294. Berlin, Germany: Springer.

18. Rujević, B., Kaplar M, Kaplar S, Stanković, R., Obradović, I., Mačutek J. 2021 Quantitative analysis of syllable properties in Croatian, Serbian, Russian, and Ukrainian. In *Current issues in linguistic theory*, vol. 356 (eds A Pawłowski, J Mačutek, S Embleton, G Mikros), pp. 56–67. Amsterdam, The Netherlands: John Benjamins Publishing Company.

19. Luo B, Jiang T, Liu Y, Wang J, Lin A, Wei X, Feng J. 2013 Brevity is prevalent in bat short-range communication. *J. Comp. Physiol. A* **199**, 325–333. (doi:10.1007/s00359-013-0793-y)

20. Favaro L, Gamba M, Cresta E, Fumagalli E, Bandoli F, Pilenga C, Isaja V, Mathevon N, Reby D. 2020 Do penguins' vocal sequences conform to linguistic laws? *Biol. Lett.* **16**, 20190589. (doi:10.1098/rsbl.2019.0589)

21. Demartsev V, Gordon N, Barocas A, Bar-Ziv E, Ilany T, Goll Y, Ilany A, Geffen E. 2019 The 'Law of Brevity' in animal communication: sex-specific signaling optimization is determined by call amplitude rather than duration. *Evol. Lett.* **3**, 623–634. (doi:10.1002/evl3.147)

22. Semple S, Hsu MJ, Agoramoorthy G, Ferrer-i-Cancho R. 2013 The law of brevity in macaque vocal communication is not an artefact of analysing mean call durations. *J. Quant. Linguist.* **20**, 209–217. (doi:10.1080/09296174.2013.799917)

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

24. Huang M, Ma H, Ma C, Garber PA, Fan P. 2020 Male gibbon loud morning calls conform to Zipf's law of brevity and Menzerath's law: insights into the origin of human language. *Anim. Behav.* **160**, 145–155. (doi:10.1016/j.anbehav.2019.11.017)

25. Valente D et al. 2021 Linguistic laws of brevity: conformity in *Indri indri*. *Anim. Cogn.* **24**, 897–906. (doi:10.1007/s10071-021-01495-3)

26. Altmann G. 1980 Prolegomena to Menzerath's law. *Glottometrika* **2**, 1–10.

27. Köhler R. 2012 *Quantitative syntax analysis*. Berlin, Germany: De Gruyter Mouton.

28. Andres J, Benešová M, Langer J. 2021 Towards a fractal analysis of the sign language. *J. Quant. Linguist.* **28**, 77–94. (doi:10.1080/09296174.2019.1656149)

29. Ferrer-i-Cancho R, Forns N. 2009 The self-organization of genomes. *Complexity* **15**, 34–36. (doi:10.1002/cplx.20296)

30. Li W. 2012 Menzerath's law at the gene-exon level in the human genome. *Complexity* **17**, 49–53. (doi:10.1002/cplx.20398)

31. Boroda MG, Altmann G. 1991 Menzerath's law in musical texts. *Musikometrica* **3**, 1–13.

32. Gustison ML, Semple S, Ferrer-i-Cancho R, Bergman TJ. 2016 Gelada vocal sequences follow Menzerath's linguistic law. *Proc. Natl Acad. Sci. USA* **113**, E2750–E2758. (doi:10.1073/pnas.1522072113)

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

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

35. Clink DJ, Ahmad AH, Klinck H. 2020 Brevity is not a universal in animal communication: evidence for compression depends on the unit of analysis in small ape vocalizations. *R. Soc. Open Sci.* **7**, 200151. (doi:10.1098/rsos.200151)

36. Watson SK, Heesen R, Hedwig D, Robbins MM, Townsend SW. 2020 An exploration of Menzerath's law in wild mountain gorilla vocal

sequences. *Biol. Lett.* **16**, 20200380. (doi:10.1098/rsbl.2020.0380)

37. Ferrer-i-Cancho R, Lusseau D. 2009 Efficient coding in dolphin surface behavioral patterns. *Complexity* **14**, 23–25. (doi:10.1002/cplx.20266)

38. Byrne RW, Cartmill E, Genty E, Graham KE, Hobaiter C, Tanner J. 2017 Great ape gestures: intentional communication with a rich set of innate signals. *Anim. Cogn.* **20**, 755–769. (doi:10.1007/s10071-017-1096-4)

39. Bard KA, Maguire-Herring V, Tomonaga M, Matsuzawa T. 2019 The gesture 'Touch': does meaning-making develop in chimpanzees' use of a very flexible gesture? *Anim. Cogn.* **22**, 535–550. (doi:10.1007/s10071-017-1136-0)

40. Call J, Tomasello M. (eds). 2007 *The gestural communication of apes and monkeys*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

41. Hobaiter C, Byrne RW. 2011 The gestural repertoire of the wild chimpanzee. *Anim. Cogn.* **14**, 745–767. (doi:10.1007/s10071-011-0409-2)

42. Liebal K, Call J, Tomasello M, Pika S. 2004 To move or not to move: how apes adjust to the attentional state of others. *Interact. Stud., Soc. Behav. Commun. Biol. Artif. Syst.* **5**, 199–219. (doi:10.1075/is.5.2.03lie)

43. Graham KE, Hobaiter C, Ounsley J, Furuichi T, Byrne RW. 2018 Bonobo and chimpanzee gestures overlap extensively in meaning. *PLoS Biol.* **16**, e2004825. (doi:10.1371/journal.pbio.2004825)

44. Hobaiter C, Byrne RW. 2014 The meanings of chimpanzee gestures. *Curr. Biol.* **24**, 1596–1600. (doi:10.1016/j.cub.2014.05.066)

45. Schel AM, Townsend SW, Machanda Z, Zuberbühler K, Slocombe KE. 2013 Chimpanzee alarm call production meets key criteria for intentionality. *PLoS ONE* **8**, e76674. (doi:10.1371/journal.pone.0076674)

46. Miton H, Morin O. 2019 When iconicity stands in the way of abbreviation: no Zipfian effect for figurative signals. *PLoS ONE* **14**, e0220793. (doi:10.1371/journal.pone.0220793)

47. Chaabouni R, Kharitonov E, Dupoux E, Baroni M. 2019 Anti-efficient encoding in emergent communication. In *Advances in Neural Information Processing Systems 32: 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada* (eds H Wallach, H Larochelle, A Beygelzimer, F D'Alché-Buc, E Fox, R Garnett), pp. 6261-6271. San Diago, CA: Neural Information Processing Systems Foundation Inc.

48. Held SDE, Špinka M. 2011 Animal play and animal welfare. *Anim. Behav.* **81**, 891–899. (doi:10.1016/j.anbehav.2011.01.007)

49. Pellis SM, Pellis VC. 1996 On knowing it's only play: the role of play signals in play fighting. *Aggress. Violent Behav.* **1**, 249–268. (doi:10.1016/1359-1789(95)00016-X)

50. Smith E. 1978 *Social play in primates*. London, UK: Academic Press.

51. Kendon A. 2004 *Gesture: visible action as utterance*. Cambridge, UK: Cambridge University Press.

52. Hatzigeorgiu N, Mikros G, Carayannis G. 2001 Word length, word frequencies and Zipf's law in the greek language. *J. Quant. Linguist.* **8**, 175–185. (doi:10.1076/jqul.8.3.175.4096)

13

53. Hobaiter C, Byrne RW. 2012 Gesture use in consortship: wild chimpanzees' use of gesture for an 'evolutionarily urgent' purpose. In *Gesture studies*, vol. 6 (eds S Pika, K Liebal), pp. 129–146. Amsterdam, The Netherlands: John Benjamins Publishing Company.

54. Nishida T. 1980 The leaf-clipping display: a newly-discovered expressive gesture in wild chimpanzees. *J. Hum. Evol.* **9**, 117–128. (doi:10.1016/0047-2484(80)90068-8)

55. Tutin CEG. 1979 Mating patterns and reproductive strategies in a community of wild chimpanzees (*Pan troglodytes schweinfurthii*). *Behav. Ecol. Sociobiol.* **6**, 29–38. (doi:10.1007/BF00293242)

56. Watts DP. 2015 Mating behavior of adolescent male chimpanzees (*Pan troglodytes*) at Ngogo, Kibale National Park, Uganda. *Primates* **56**, 163–172. (doi:10.1007/s10329-014-0453-z)

57. Clark CB. 1977 A preliminary report on weaning among chimpanzees of the Gombe National Park, Tanzania. In *Primate bio-social development: biological, social and ecological determinants* (eds S Chevalier-Skolnikoff, FE Poirier), pp. 235–260. New York, NY: Garland.

58. Thompson ME. 2013 Reproductive ecology of female chimpanzees: chimpanzee reproductive ecology. *Am. J. Primatol.* **75**, 222–237. (doi:10.1002/ajp.22084)

59. Newton-Fisher NE, Thompson ME, Reynolds V, Boesch C, Vigilant L. 2009 Paternity and social rank in wild chimpanzees (*Pan troglodytes*) from the Budongo Forest, Uganda. *Am. J. Phys. Anthropol.* **142**, 417–428. (doi:10.1002/ajpa.21241)

60. Fawcett K, Muhumuza G. 2000 Death of a wild chimpanzee community member: possible outcome of intense sexual competition. *Am. J. Primatol.* **51**, 243–247. (doi:10.1002/1098-2345(200008)51:4<243::AID-AJP3>3.0.CO;2-P)

61. Wilson ML *et al.* 2014 Lethal aggression in *Pan* is better explained by adaptive strategies than human impacts. *Nature* **513**, 414–417. (doi:10.1038/nature13727)

62. Hobaiter C, Byrne RW, Zuberbühler K. 2017 Wild chimpanzees' use of single and combined vocal and gestural signals. *Behav. Ecol. Sociobiol.* **71**, 96. (doi:10.1007/s00265-017-2325-1)

63. Hobaiter C, Byrne RW. 2011 Serial gesturing by wild chimpanzees: its nature and function for communication. *Anim. Cogn.* **14**, 827–838. (doi:10.1007/s10071-011-0416-3)

64. Muller MN, Wrangham RW. (Eds.). 2009 *Sexual coercion in primates and humans: an evolutionary perspective on male aggression against females*. Cambridge, MA: Harvard University Press.

65. Scott CB. 2008 *A primer for the exercise and nutrition sciences: thermodynamics, bioenergetics, metabolism*. Berlin, Germany: Springer.

66. Graham KE, Badihi G, Safryghin A, Grund C, Hobaiter C. 2022 A socio-ecological perspective on the gestural communication of great ape species, individuals, and social units. *Ethol. Ecol. Evol.* **34**, 235–259. (doi:10.1080/03949370.2021.1988722)

67. Kanwal J, Smith K, Culbertson J, Kirby S. 2017 Zipf's law of abbreviation and the principle of least effort: language users optimise a miniature lexicon for efficient communication. *Cognition* **165**, 45–52. (doi:10.1016/j.cognition.2017.05.001)

68. Janik VM, Slater PJB. 1997 Vocal learning in mammals. *Adv. Study Behav.* **26**, 59–99.

69. Fitch WT, de Boer B, Mathur N, Ghazanfar AA. 2016 Monkey vocal tracts are speech-ready. *Sci. Adv.* **2**, e1600723. (doi:10.1126/sciadv.1600723)

70. Altmann J. 1974. Observational study of behavior: sampling methods. *Behavior* **49**, 227–266.

71. Landers R. 2015 Computing intraclass correlations (ICC) as estimates of interrater reliability in SPSS (dataset). *Winnower* **2**, e143518.

72. R Core Team. 2020 *R: a language and environment for statistical computing (4.0.0)*. Vienna, Austria: R Foundation for Statistical Computing. See https://www.R-project.org/.

73. 2020 RStudio: integrated development environment for R (1.2.5042). See http://www.rstudio.com/.

74. Bürkner P-C. 2017 brms: an R package for Bayesian multilevel models using *Stan*. *J. Stat. Softw.* **80**. (doi:10.18637/jss.v080.i01)

75. Rodrigues ED, Santos AJ, Veppo F, Pereira J, Hobaiter C. 2021 Connecting primate gesture to the evolutionary roots of language: a systematic review. *Am. J. Primatol.* **83**, e23313 (doi:10.1002/ajp.23313)

76. Vehtari A, Gelman A, Gabry J. 2017 Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat. Comput.* **27**, 1413–1432. (doi:10.1007/s11222-016-9696-4)

77. 2022 'RStan: the R interface to stan.' (version 2.21.5). See https://mc-stan.org/.

78. Delignette-Muller ML, Dutang C. 2015 {fitdistrplus}: an {R} package for fitting distributions. *J. Stat. Softw.* **64**, 1–34. (doi:10.18637/jss.v064.i04)

79. Safryghin A, Cross C, Fallon B, Heesen R, Ferrer-i-Cancho R, Hobaiter C. 2022 Variable expression of linguistic laws in ape gesture: a case study from chimpanzee sexual solicitation. Figshare. (doi:10.6084/m9.figshare.c.6268394)