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

Short title: Linguistic laws in chimpanzee solicitation gestures.

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

[1]School of Psychology and Neuroscience, University of St Andrews, St Andrews, Fife, United Kingdom

[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, United Kingdom

[4]Budongo Conservation Field Station, Masindi, Uganda

* Corresponding author

Email: clh42@st-andrews.ac.uk (CH)

[&] These authors contributed equally to this work

Keywords: compression, communication, Zipf, Menzerath, language

Abstract

Two language laws have been identified as consistent patterns shaping animal behaviour, both acting on the organisational 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 minimise 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 that, in addition to selection on the signal repertoire, ape gestural expression appears shaped by factors in the immediate socio-ecological context.

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 (Altmann & Gerlach, 2016; Bentz & Ferrer-I-Cancho, 2016; Börstell et al., 2016; Hernández-Fernández et al., 2019; Köhler et al., 2005; Menzerath, 1954; Naranan & Balasubrahmanyan, 2000; Sanada, 2008; Semple et al., 2022; Wang & Chen, 2015; Zipf, 1936). These regularities are hypothesized to be manifestations of the information theoretic principle of compression (Ferrer-i-Cancho, Bentz, et al., 2022; Semple et al., 2022). Compression is a particular case of the principle of least effort (Zipf, 1949) – a principle that promotes the outcome that requires the least amount of energy to produce or achieve – and thereby promotes coding efficiency (Ferrer-i-Cancho et al., 2013). In communication, compression is expressed as a pressure towards reducing the energy needed to compose a code but limited by the need to retain the critical information in the transmission (Cover & Thomas, 2006; Ferrer-i-Cancho et al., 2022).

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 (Strauss et al., 2007; Zipf, 1949), and is generalised as the tendency for more frequent elements of many kinds (*e.g.*, syllables, words, calls) to be shorter or smaller (Ferrer-i-Cancho et al., 2013) – with similar patterns found at different levels of analysis, for example in speech at the level of words (Strauss et al., 2007) syllables (Rujević et al., 2021), and phonemes (Hernández-Fernández et al., 2019). As well as being found in human spoken, signed, and written languages (Bentz & Ferrer-I-Cancho, 2016; Börstell et al., 2016; Hernández-Fernández et al., 2019; Sanada, 2008; Wang & Chen, 2015), Zipf's law of brevity has been identified in the short-range communication of diverse taxa: dolphins (Ferrer-i-Cancho et al., 2022), bats (Luo et al., 2013), penguins (Favaro et al., 2020), hyraxes (Demartsev et al., 2019), and various primates (macaques: Semple et al., 2013; marmosets: Ferrer-i-Cancho & Hernández-Fernández, 2013; gibbons: Huang et al., 2020; Indri indri: Valente et al., 2021), as well as in genomes (Naranan & Balasubrahmanyan, 2000).

At the level of constructs, Menzerath's law states that "*the greater the whole, the smaller its constituents*" (Altmann, 1980; Köhler, 2012; Menzerath, 1954); 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 (Altmann, 1980; Andres et al., 2021), genomes (Ferrer-i-Cancho & Forns, 2009; Li, 2012), music (Boroda & Altmann, 1991), and in the communication of dolphins (Ferrer-i-Cancho et al., 2022), penguins (Favaro et al., 2020), and primates (geladas: Gustison et al., 2016; chimpanzees: Fedurek et al., 2017; Heesen et al., 2019; gibbons: Clink et al., 2020; Huang et al., 2020; gorillas: Watson et al., 2020; Indri indri: Valente et al., 2021;). 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 (Börstell et al., 2016). Zipf's law of brevity was also found in fingerspelling, with a negative relationship between mean fingerspelled sign duration and frequency (Börstell et al., 2016). Similarly, Czech sign language was found to follow Menzerath's law (Andres et al., 2021). 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; Ferrer-i-Cancho & Lusseau, 2009) 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 (Heesen et al., 2019).

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 (Byrne et al., 2017); as compared to vocal communication, gestural repertoires are larger and are more flexibly deployed, with individual gesture types used to achieve multiple goals (Bard et al., 2019; Call & Tomasello, 2007; Hobaiter & Byrne, 2011a; Liebal et al., 2004). Gestures are also used intentionally, *i.e.*, to reach social goals by influencing the receivers' behaviour or understanding (Graham et al., 2018; Hobaiter & Byrne, 2011a, 2014; Schel et al., 2013), and flexibly across contexts (Call & Tomasello, 2007; Hobaiter & Byrne, 2011a; Liebal et al., 2004). Nevertheless, Heesen et al.'s (2019) 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 (Semple et al., 2022).

Although a lack of evidence supporting Zipf's law of brevity has been previously reported (e.g., European heraldry: Miton & Morin, 2019; computer-based neural-networks: Chaabouni et al., 2019), these remain rare exceptions, and in non-human animal communication have typically only been reported in long-distance vocal communication (*e.g.* gibbon song: Clink et al., 2020; bats: Luo et al., 2013; although *cf*. female hyrax calls: Demartsev et al., 2019) where the impact of distance on signal transmission fidelity may have a particularly strong effect on the costs of compression (Ferrer-i-Cancho et al., 2013;

Gustison et al., 2016; Semple et al., 2022). 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 (Semple et al., 2022). Play is produced when there is an excess of time and energy (Held & Špinka, 2011; Pellis & Pellis, 1996; Smith, 2014), 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 due 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 (Kendon, 2004). 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. In 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., Hatzigeorgiu et al., 2001). In 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) (Hobaiter & Byrne, 2012; *c.f.* Call & Tomasello, 2007). 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 (Hobaiter & Byrne, 2012; Nishida, 1980). While a range of gesture types are employed, these are typically a smaller sub-set of the available repertoire – *c.f.* 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 (Tutin, 1979). Both direct solicitation and consortship and are key strategies for individual fitness (Tutin, 1979; Watts, 2015), and as such behaviour associated with them is likely subject to strong selective pressures. The energetic costs of lactation mean that adult female chimpanzees typically concieve only once every 4-5 years (Clark, 1977; Thompson, 2013). So while there are typically 60-80 individuals in a group, the operational sex ratio of available females in estrus may be very small, and males show substantial variation in reproductive success (Newton-Fisher et al., 2009; Tutin, 1979). 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 (Fawcett & Muhumuza, 2000; Tutin, 1979) as well as potentially lethal attacks from males in neighbouring groups (Wilson et al., 2014). 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 (Hobaiter et al., 2017). Therefore, on one hand individuals benefit from producing conspicuous energetic signals to attract females, often having to insist to secure mating; on 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 generalised 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 vs whole body) as fixed factors, and signaller's 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**.**

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 5 whole-body gestures (Figure 1; for definitions for full repertoire definitions see Table S1 in supporting Information 1) performed by 16 male chimpanzees aged 10-42 years old. On average, each 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 supplementary information.

Gesture token duration was measured via analysis of video data with a minimum unit of 0.04s (one frame). Duration ranged from 0.04-15.04 seconds (median: 1.56 ±2.35s). If consecutive gesture tokens were performed with less than 1s in between them, they were considered to form a sequence (Heesen et al., 2019; Hobaiter & Byrne, 2011b). 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. 244 sequences were composed of a single token, the remaining 115 sequences had length *n*>1. Of the 115 sequences analysed that were composed of 2 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).

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

To test for Zipf's law we ran a Bayesian generalised 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 Supporting information 2 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 ± 0.3). For Zipf-model effects Bulk 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 (Supporting information 3, Table S5; b = 0.90, s.d. = 1.26, 95% Credible Intervals (CrI) [−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; Supporting information 3, 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; Supporting information 3, Table S7).

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 proportion of whole-body gestures within the sequence (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 (Supporting information 3, Table S8; b = -0.18, s.d. = 0.04, 95% CrI [-0.26, -

0.11]; Figure 3A). Similar results were found when 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 (Supporting information 3, Table S9, Figure 3B; b = -0.23, s.d. = 0.04, 95% CrI [-0.31, -0.15]). In 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; Supporting information 3; 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 (Supporting Information 4).

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 3A and Figure 3B. In addition, longer sequences were formed of a) a mix of *loose* and *fixed* duration gestures or b) only *loose* duration gestures (see supporting information 5, 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.

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 (*c.f.* Heesen et al., 2019 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 the Duane's data and in the full dataset, as well as in a range of alternative analyses (Supporting Information 6). 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 data set.

These results represent a further absence of evidence in support of Zipf's law of brevity in great ape gestural communication (Heesen et al., 2019) 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 (Börstell et al., 2016; Ferrer-i-Cancho et al., 2013; Semple et al., 2022). 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 (Heesen et al., 2019). 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 (Supporting Information 6). Visual inspection of the 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 randomised subsetting of sufficiently large datasets. As Semple et al. (2022) 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.

In 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 (Muller & Wrangham, 2009), chimpanzee paternity is often heavily biased towards higher-ranking individuals (Newton-Fisher et al., 2009). 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 (Hobaiter et al., 2017). While there are examples of vocal communication systems used in biologically ‘relevant’ contexts that adhere to Zipf’s brevity law (Favaro et al., 2020), 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 these prolonged and conspicuous signals appear to remain constrained by physiological mechanisms of gestural production. As for primate vocal communication (Fedurek et al., 2017; Gustison et al., 2016), 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 (Scott, 2008) 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 vocalisations, that is not the case for female vocalisations where Zipf's law of brevity emerges when analysing call amplitude rather than duration (Demartsev et al., 2019). Conversely, in Börstell et al. (2016) research on Swedish Sign Language, Zipf's law of brevity seems to hold across sign categorisation, 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 (Byrne et al., 2017; Graham et al., 2022). 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. In 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 (Börstell et al., 2016). 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 likely 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 (Fawcett & Muhumuza, 2000). 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 (Kanwal et al., 2017). Conversely, when participants were required to produce solely time-efficient vs 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 (Janik & Slater, 1997; Fitch et al., 2016), gestural flexibility (in goal and context – Bard et al., 2019; Call & Tomasello, 2007; Hobaiter & Byrne, 2011a; Liebal et al., 2004) 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 (Graham et al., 2022). 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.

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 7.30am and 4.30pm with recording of gestures following a focal behaviour sampling approach (Altmann, 1974). 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) were used 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.

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 (Hobaiter & Byrne, 2011a, 2012). 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 (Tutin, 1979). 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.

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 S4 Table 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 (Hobaiter & Byrne, 2011b): 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 second 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 1s between two consecutive gesture tokens; single gestures were coded as sequences of length one (Heesen et al., 2019; Hobaiter & Byrne, 2011b).

Gesture duration

Gesture duration was calculated using MPEG streamclip (version 1.9.3beta). We measured gesture duration in frames, each lasting 0.04s. 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 a preparation, action stroke, hold or repetition, and recovery phase (Kendon, 2004). Here we follow previous work in (Heesen et al., 2019) 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 (1) the cessation of the body movement related to gesture production, or (2) a change in body positioning if the gesture relied on body alignment, or (3) 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 (1) a change in gesture action, e.g., from a reach to a shake, (2) a change in the rhythm or orientation of a gesture action, hold, or repetition, e.g., the rhythm or direction of an object shake is broken or changed (Hobaiter & Byrne, 2017).

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 (Landers, 2015) – which revealed very high agreement on gesture duration measurements (ICC=0.995, p<.001). Unfortunately, an additional step of inter-observer reliability was not possible due to the loss of the file that linked the original dataset to the videos from which data were extracted.

Statistical analysis

All data were analysed using R version 4.0.0 and RStudio version 1.2.5042 (R Core Team, 2020; RStudio Team, 2020).We fitted Bayesian generalised linear multivariate multilevel models using the 'brm' function from the 'brms' package (Bürkner, 2017) with minimally informative priors, 2000 iterations and 3 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 vs 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 (Heesen et al., 2019; Rodrigues et al., 2021) or in signed languages and fingerspelling (e.g., Börstell et al., 2016).

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 proportion of whole-body gestures within the sequence (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 Level One Out information criterion (LOO) (Vehtari et al., 2017) 'loo_compare' function from the 'stan' package (version 2.21.5; Stan Development Team, 2022) 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 (version 1.0-14; Delignette-Muller & Dutang, 2015). 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 supporting information 2).

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 generalised 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 supporting information 6 to allow for comparison with previous work that analysed median durations with or without implementing generalised linear models (e.g., Hernández-Fernández et al. 2019; Watson et al. 2020).

Data and code

Data and code for all analyses are available in a public GitHub repository: github.com/Wild-Minds/LinguisticLaws_Papers

Acknowledgements

We thank the staff and field assistants of the Budongo Conservation Field Station for their assistance in the original gestural data collection, and the Ugandan National Council for Science and Technology and the Ugandan Wildlife Authority for permission to conduct the original research. We thank the Royal Zoological Society of Scotland for its funding of the field station. We thank Dr Alexander Mielke for his advice on the statistical models. We thank the editor and three anonymous reviewers for their constructive comments. This research received funding from the European Union's 8th Framework Programme, Horizon 2020, under grant agreement no 802719.

References

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

535 Altmann, G., & Gerlach, M. (2016). Statistical Laws in Linguistics. In M. D. Esposti, Altmann, & F. Pachet (Eds.), *Creativity and Universality in Language* (pp. 7–26). Springer International Publishing. https://doi.org/10.1007/978-3-319-24403-7_2

538 Andres, J., Benešová, M., & Langer, J. (2021). Towards a fractal analysis of the sign language. *Journal of Quantitative Linguistics*, *28*(1), 77–94. https://doi.org/10.1080/09296174.2019.1656149

541 Bard, K. A., Maguire-Herring, V., Tomonaga, M., & Matsuzawa, T. (2019). The gesture 'Touch': Does meaning-making develop in chimpanzees' use of a very flexible gesture? *Animal Cognition*, *22*(4), 535–550. https://doi.org/10.1007/s10071-017-1136-0

545 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, Netherlands. https://doi.org/10.15496/PUBLIKATION-10057

548 Boroda, M. G., & Altmann, G. (1991). Menzerath's law in musical texts. *Musikometrica*, *3*, 1–13.

550 Börstell, C., Hörberg, T., & Östling, R. (2016). Distribution and duration of signs and parts of speech in Swedish Sign Language. *Sign Language & Linguistics*, *19*(2), 143–196. https://doi.org/10.1075/sll.19.2.01bor

553 Bürkner, P.-C. (2017). **brms**: An *R* Package for Bayesian Multilevel Models Using *Stan*. *Journal of Statistical Software*, *80*(1). https://doi.org/10.18637/jss.v080.i01

555 Byrne, R. W., Cartmill, E., Genty, E., Graham, K. E., Hobaiter, C., & Tanner, J. (2017). Great ape gestures: Intentional communication with a rich set of innate signals. *Animal Cognition*, *20*(4), 755–769. https://doi.org/10.1007/s10071-017-1096-4

Call, J., & Tomasello, M. (Eds.). (2007). *The gestural communication of apes and monkeys*. Lawrence Erlbaum Associates, Publishers.

Chaabouni, R., Kharitonov, E., Dupoux, E., & Baroni, M. (2019). *Anti-efficient encoding in emergent communication.* 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada.

Clark, C. B. (1977). A preliminary report on weaning among chimpanzees of the Gombe National Park, Tanzania. In S. Chevalier-Skolnikoff & F. E. Poirier (Eds.), *Primate Bio-social Development: Biological, Social and Ecological Determinants.* (pp. 235–260). Garland.

Clink, D. J., Ahmad, A. H., & 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. https://doi.org/10.1098/rsos.200151

Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory* (2nd ed). Wiley-Interscience.

Delignette-Muller, M. L., & Dutang, C. (2015). {fitdistrplus}: An {R} Package for Fitting Distributions. *Journal of Statistical Software*, *64*(4), 1–34.

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. *Evolution Letters*, *3*(6), 623–634. https://doi.org/10.1002/evl3.147

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? *Biology Letters*, *16*(2), 20190589. https://doi.org/10.1098/rsbl.2019.0589

Fawcett, K., & Muhumuza, G. (2000). Death of a wild chimpanzee community member: Possible outcome of intense sexual competition. *American Journal of Primatology*, *51*(4), 243–247. https://doi.org/10.1002/1098-2345(200008)51:4<243::AID-AJP3>3.0.CO;2-P

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. *Frontiers in Zoology*, *14*(1), 50. https://doi.org/10.1186/s12983-017-0235-8

Ferrer-i-Cancho, R., Bentz, C., & Seguin, C. (2022). Optimal coding and the origins of Zipfian laws. *Journal of Quantitative Linguistics*, *29*(2), 165–194. https://doi.org/10.1080/09296174.2020.1778387

Ferrer-i-Cancho, R., & Forns, N. (2009). The self-organization of genomes. *Complexity*, *15*(5), 34-36. https://doi.org/10.1002/cplx.20296

Ferrer-i-Cancho, R., & Hernández-Fernández, A. (2013). The failure of the law of brevity in two new world primates. Statistical Caveats. *Glottotheory*, *4*(1), 45-55. https://doi.org/10.1524/glot.2013.0004

Ferrer-i-Cancho, R., Hernández-Fernández, A., Baixeries, J., Dębowski, Ł., & Mačutek, J. (2014). When is Menzerath-Altmann law mathematically trivial? A new approach. *Statistical Applications in Genetics and Molecular Biology*, *13*(6), 633-644. https://doi.org/10.1515/sagmb-2013-0034

Ferrer-i-Cancho, R., Hernández-Fernández, A., Lusseau, D., Agoramoorthy, G., Hsu, M. J., & Semple, S. (2013). Compression as a universal principle of animal behavior. *Cognitive Science*, *37*(8), 1565–1578. https://doi.org/10.1111/cogs.12061

Ferrer-i-Cancho, R., & Lusseau, D. (2009). Efficient coding in dolphin surface behavioral patterns. *Complexity*, *14*(5), 23–25. https://doi.org/10.1002/cplx.20266

Ferrer-i-Cancho, R., Lusseau, D., & McCowan, B. (2022). Parallels of human language in the behavior of bottlenose dolphins. *Linguistic Frontiers*, 5(1), 1-7. https://doi.org/10.2478/lf-2022-0002

Fitch, W. T., de Boer, B., Mathur, N., & Ghazanfar, A. A. (2016). Monkey vocal tracts are speech-ready. *Science Advances*, *2*(12), e1600723. https://doi.org/10.1126/sciadv.1600723

Graham, K. E., 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. *Ethology Ecology & Evolution*, *34*(3), 235–259. https://doi.org/10.1080/03949370.2021.1988722

Graham, K. E., Hobaiter, C., Ounsley, J., Furuichi, T., & Byrne, R. W. (2018). Bonobo and chimpanzee gestures overlap extensively in meaning. *PLOS Biology*, *16*(2), e2004825. https://doi.org/10.1371/journal.pbio.2004825

Gustison, M. L., Semple, S., Ferrer-i-Cancho, R., & Bergman, T. J. (2016). Gelada vocal sequences follow Menzerath's linguistic law. *Proceedings of the National Academy of Sciences*, *113*(19), E2750–E2758. https://doi.org/10.1073/pnas.1522072113

Hatzigeorgiu, N., Mikros, G., & Carayannis, G. (2001). Word length, word frequencies and Zipf's law in the greek language. *Journal of Quantitative Linguistics*, *8*(3), 175–185. https://doi.org/10.1076/jqul.8.3.175.4096

Heesen, R., Hobaiter, C., Ferrer-i-Cancho, R., & Semple, S. (2019). Linguistic laws in chimpanzee gestural communication. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1896), 20182900. https://doi.org/10.1098/rspb.2018.2900

Held, S. D. E., & Špinka, M. (2011). Animal play and animal welfare. *Animal Behaviour*, *81*(5), 891–899. https://doi.org/10.1016/j.anbehav.2011.01.007

Hernández-Fernández, A., G. Torre, I., Garrido, J.-M., & Lacasa, L. (2019). Linguistic laws in speech: The case of Catalan and Spanish. *Entropy*, *21*(12), 1153. https://doi.org/10.3390/e21121153

Hobaiter, C., & Byrne, R. W. (2011a). The gestural repertoire of the wild chimpanzee. *Animal Cognition*, *14*(5), 745–767. https://doi.org/10.1007/s10071-011-0409-2

Hobaiter, C., & Byrne, R. W. (2011b). Serial gesturing by wild chimpanzees: Its nature and function for communication. *Animal Cognition*, *14*(6), 827–838. https://doi.org/10.1007/s10071-011-0416-3

Hobaiter, C., & Byrne, R. W. (2012). Gesture use in consortship: Wild chimpanzees' use of gesture for an 'evolutionarily urgent' purpose. In S. Pika & K. Liebal (Eds.), *Gesture Studies* (Vol. 6, pp. 129–146). John Benjamins Publishing Company. https://doi.org/10.1075/gs.6.08hob

Hobaiter, C., & Byrne, R. W. (2014). The meanings of chimpanzee gestures. *Current Biology*, *24*(14), 1596–1600. https://doi.org/10.1016/j.cub.2014.05.066

Hobaiter, C., Byrne, R. W., & Zuberbühler, K. (2017). Wild chimpanzees' use of single and combined vocal and gestural signals. *Behavioral Ecology and Sociobiology*, *71*(6), 96. https://doi.org/10.1007/s00265-017-2325-1

Huang, M., Ma, H., Ma, C., Garber, P. A., & 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. *Animal Behaviour*, *160*, 145–155. https://doi.org/10.1016/j.anbehav.2019.11.017

Janik, V. M., & Slater, P. J. B. (1997). Vocal Learning in Mammals. In *Advances in the Study of Behavior* (Vol. 26, pp. 59–99). Elsevier. https://doi.org/10.1016/S0065-3454(08)60377-0

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. https://doi.org/10.1016/j.cognition.2017.05.001

Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge University Press. https://doi.org/10.1017/CBO9780511807572

Köhler, R. (2012). *Quantitative syntax analysis*. De Gruyter Mouton.

Köhler, R., Altmann, G., & Piotrowski, R. G. (Eds.). (2005). *Quantitative Linguistik / Quantitative Linguistics: Ein internationales Handbuch / An International Handbook*. De Gruyter Mouton. https://doi.org/10.1515/9783110155785

Landers, R. (2015). Computing intraclass correlations (ICC) as estimates of interrater reliability in SPSS [Data set]. In *The Winnower* (Vol. 2, p. e143518). https://doi.org/10.15200/winn.143518.81744

Li, W. (2012). Menzerath's law at the gene-exon level in the human genome. *Complexity*, *17*(4), 49–53. https://doi.org/10.1002/cplx.20398

Liebal, K., Call, J., Tomasello, M., & Pika, S. (2004). To move or not to move: How apes adjust to the attentional state of others. *Interaction Studies. Social Behaviour and Communication in Biological and Artificial Systems*, *5*(2), 199–219. https://doi.org/10.1075/is.5.2.03lie

Luo, B., Jiang, T., Liu, Y., Wang, J., Lin, A., Wei, X., & Feng, J. (2013). Brevity is prevalent in bat short-range communication. *Journal of Comparative Physiology A*, *199*(4), 325–333. https://doi.org/10.1007/s00359-013-0793-y

Menzerath, P. (1954). *Die Architektonik des deutschen Wortschatzes: Vol. Volume 3*. F. Dümmler.

Miton, H., & Morin, O. (2019). When iconicity stands in the way of abbreviation: No Zipfian effect for figurative signals. *PLOS ONE*, *14*(8), e0220793. https://doi.org/10.1371/journal.pone.0220793

Muller, M. N., & Wrangham, R. W. (Eds.). (2009). *Sexual coercion in primates and humans: An evolutionary perspective on male aggression against females*. Harvard University Press.

Naranan, S., & Balasubrahmanyan, V. K. (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. *Journal of Quantitative Linguistics*, *7*(2), 129–151. https://doi.org/10.1076/0929-6174(200008)07:02;1-Z;FT129

Newton-Fisher, N. E., Thompson, M. E., Reynolds, V., Boesch, C., & Vigilant, L. (2009). Paternity and social rank in wild chimpanzees (Pan troglodytes) from the Budongo Forest, Uganda. *American Journal of Physical Anthropology*, *142*(3), 417–428. https://doi.org/10.1002/ajpa.21241

Nishida, T. (1980). The leaf-clipping display: A newly-discovered expressive gesture in wild chimpanzees. *Journal of Human Evolution*, *9*(2), 117–128. https://doi.org/10.1016/0047-2484(80)90068-8

Nishida, T. (Ed.). (2010). *Chimpanzee behavior in the wild: An audio-visual encyclopedia*. Springer.

Pellis, S. M., & Pellis, V. C. (1996). On knowing it's only play: The role of play signals in play fighting. *Aggression and Violent Behavior*, *1*(3), 249–268. https://doi.org/10.1016/1359-1789(95)00016-X

R Core Team. (2020). *R: A Language and Environment for Statistical Computing* (4.0.0) [R]. R Foundation for Statistical Computing. https://www.R-project.org/

Rodrigues, E. D., Santos, A. J., Veppo, F., Pereira, J., & Hobaiter, C. (2021). Connecting primate gesture to the evolutionary roots of language: A systematic review. *American Journal of Primatology*, *83*(9). https://doi.org/10.1002/ajp.23313

RStudio Team. (2020). *RStudio: Integrated Development Environment for R* (1.2.5042) [R]. RStudio, Inc. http://www.rstudio.com/

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 A. Pawłowski, J. Mačutek, S. Embleton, & G. Mikros (Eds.), *Current Issues in Linguistic Theory* (Vol. 356, pp. 56–67). John Benjamins Publishing Company. https://doi.org/10.1075/cilt.356.04ruj

Sanada, H. (2008). *Investigations in Japanese historical lexicology* (Rev. ed). Peust & Gutschmidt.

Schel, A. M., Townsend, S. W., Machanda, Z., Zuberbühler, K., & Slocombe, K. E. (2013). Chimpanzee alarm call Production meets key criteria for intentionality. *PLoS ONE*, *8*(10), e76674. https://doi.org/10.1371/journal.pone.0076674

Scott, C. B. (2008). *A primer for the exercise and nutrition sciences: Thermodynamics, bioenergetics, metabolism*. Springer.

Semple, S., Ferrer-i-Cancho, R., & Gustison, M. L. (2022). Linguistic laws in biology. *Trends in Ecology & Evolution*, *37*(1), 53–66. https://doi.org/10.1016/j.tree.2021.08.012

Semple, S., Hsu, M. J., Agoramoorthy, G., & Ferrer-i-Cancho, R. (2013). The law of brevity in macaque vocal communication is not an artefact of analysing mean call durations.

Journal of Quantitative Linguistics, *20*(3), 209–217. https://doi.org/10.1080/09296174.2013.799917

Smith, E. (2014). *Social Play in Primates.* Elsevier Science. http://qut.eblib.com.au/patron/FullRecord.aspx?p=1172757

Sokal, R. R., & Rohlf, F. J. (1995). *Biometry: The principles and practice of statistics in biological research* (3rd ed). W.H. Freeman.

Stan Development Team. (2022). *“RStan: The R interface to Stan.”* (2.21.5) [R]. https://mc-stan.org/.

Strauss, U., Grzybek, P., & Altmann, G. (2007). Word Length and Word Frequency. In P. Grzybek (Ed.), *Contributions to the Science of Text and Language* (Vol. 31, pp. 277–294). Springer Netherlands. https://doi.org/10.1007/978-1-4020-4068-9_13

Thompson, M. E. (2013). Reproductive ecology of female chimpanzees: Chimpanzee reproductive ecology. *American Journal of Primatology*, *75*(3), 222–237. https://doi.org/10.1002/ajp.22084

Tutin, C. E. G. (1979). Mating patterns and reproductive strategies in a community of wild chimpanzees (Pan troglodytes schweinfurthii). *Behavioral Ecology and Sociobiology*, *6*(1), 29–38. https://doi.org/10.1007/BF00293242

Valente, D., De Gregorio, C., Favaro, L., Friard, O., Miaretsoa, L., Raimondi, T., Ratsimbazafy, J., Torti, V., Zanoli, A., Giacoma, C., & Gamba, M. (2021). Linguistic laws of brevity: Conformity in Indri indri. *Animal Cognition*, *24*(4), 897–906. https://doi.org/10.1007/s10071-021-01495-3

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413–1432. https://doi.org/10.1007/s11222-016-9696-4

Wang, Y., & Chen, X. (2015). Structural Complexity of Simplified Chinese Characters. In A. Tuzzi, M. Benesová, & J. Macutek (Eds.), *Recent Contributions to Quantitative Linguistics*. DE GRUYTER. https://doi.org/10.1515/9783110420296-019

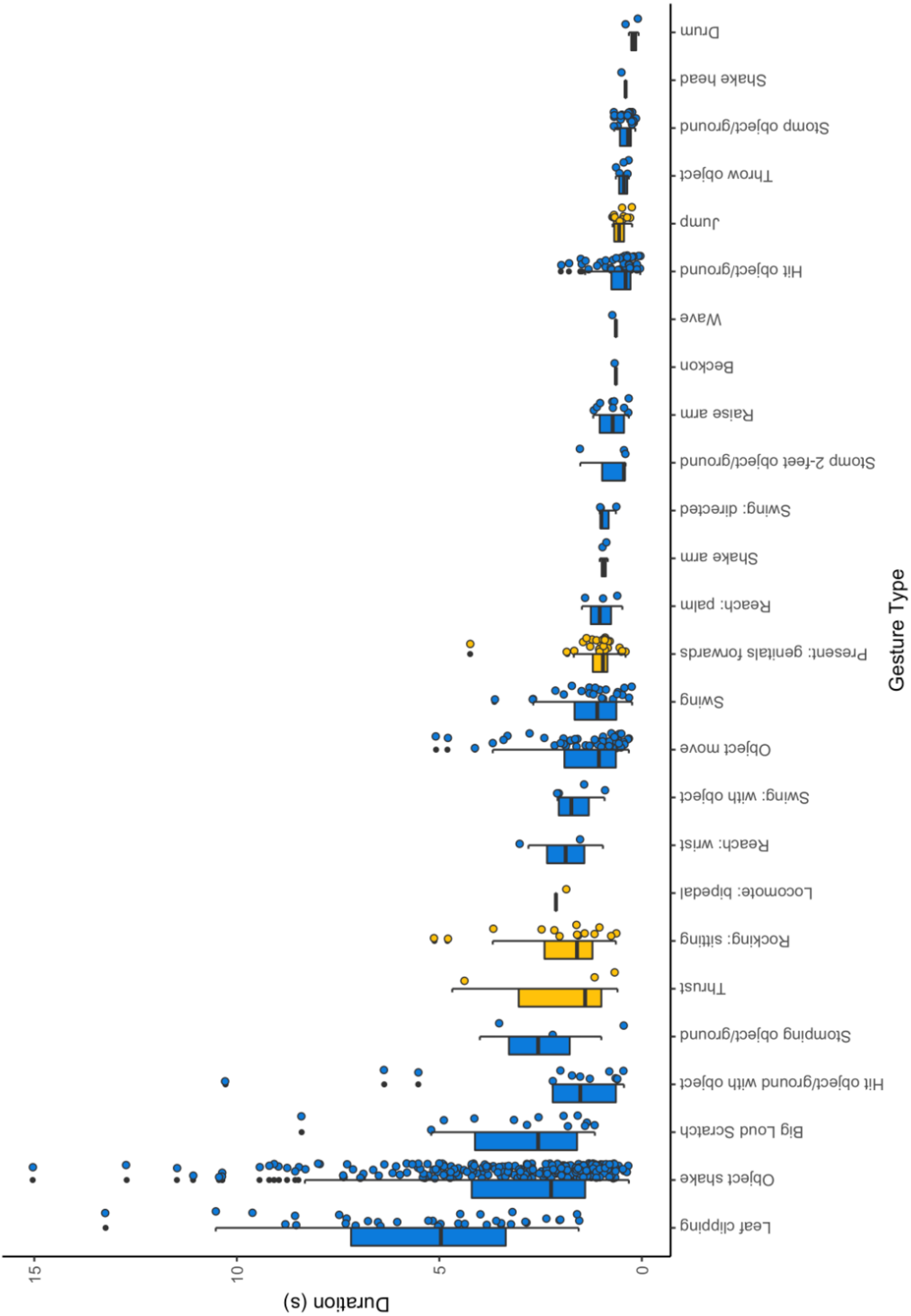
Watson, S. K., Heesen, R., Hedwig, D., Robbins, M. M., & Townsend, S. W. (2020). An exploration of Menzerath's law in wild mountain gorilla vocal sequences. *Biology Letters*, *16*(10), 20200380. https://doi.org/10.1098/rsbl.2020.0380

Watts, D. P. (2015). Mating behavior of adolescent male chimpanzees (Pan troglodytes) at Ngogo, Kibale National Park, Uganda. *Primates*, *56*(2), 163–172. https://doi.org/10.1007/s10329-014-0453-z

Wilson, M. L., Boesch, C., Fruth, B., Furuichi, T., Gilby, I. C., Hashimoto, C., Hobaiter, C. L., Hohmann, G., Itoh, N., Koops, K., Lloyd, J. N., Matsuzawa, T., Mitani, J. C., Mjungu, D. C., Morgan, D., Muller, M. N., Mundry, R., Nakamura, M., Pruetz, J., … Wrangham, R. W. (2014). Lethal aggression in Pan is better explained by adaptive strategies than human impacts. *Nature*, *513*(7518), 414–417. https://doi.org/10.1038/nature13727

Zipf, G. K. (1936). *The psycho-biology of language: An introduction to dynamic philology*. Routledge.

Zipf, G. K. (1949). *Human behavior and the principle of least effort: An introduction to human ecology*. Addison-Wesley Press.


Drum
Shake head
Stomp object/ground
Throw object
Jump
Hit object/ground
Wave
Beckon
Raise arm
Stomp 2-feet object/ground
Swing: directed
Shake arm
Reach: palm
Present: genitals forwards
Swing
Object move
Swing: with object
Reach: wrist
Locomote: bipedal
Rocking: sitting
Thrust
Stomping object/ground
Hit object/ground with object
Big Loud Scratch
Object shake
Leaf clipping
Gesture Type
15
10
5
0
Duration (s)

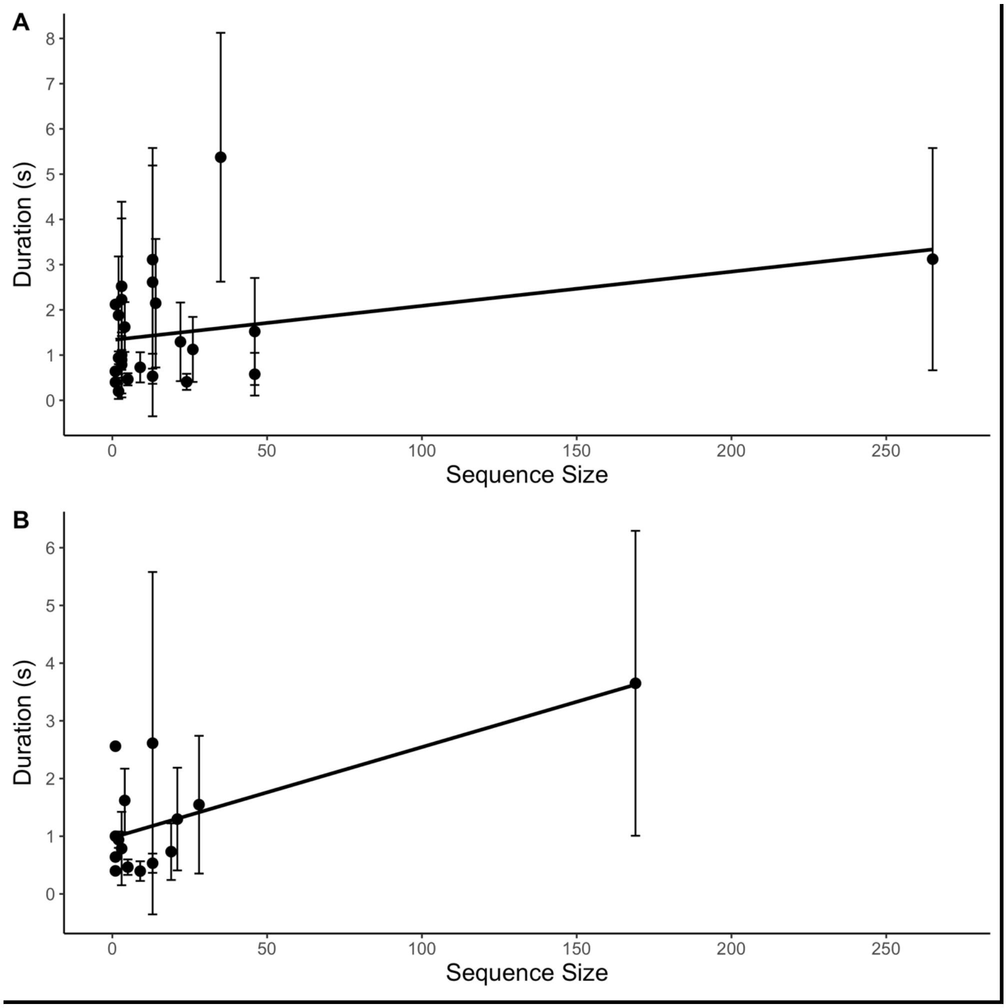


Figure 2. Relationship between frequency of occurrence and gesture duration for the full dataset (A) and Duane only data (B). Points represent the mean duration of each gesture type, with error bars showing the standard deviation from the mean. Black line indicates regression slope.

159x159mm (220 x 220 DPI)

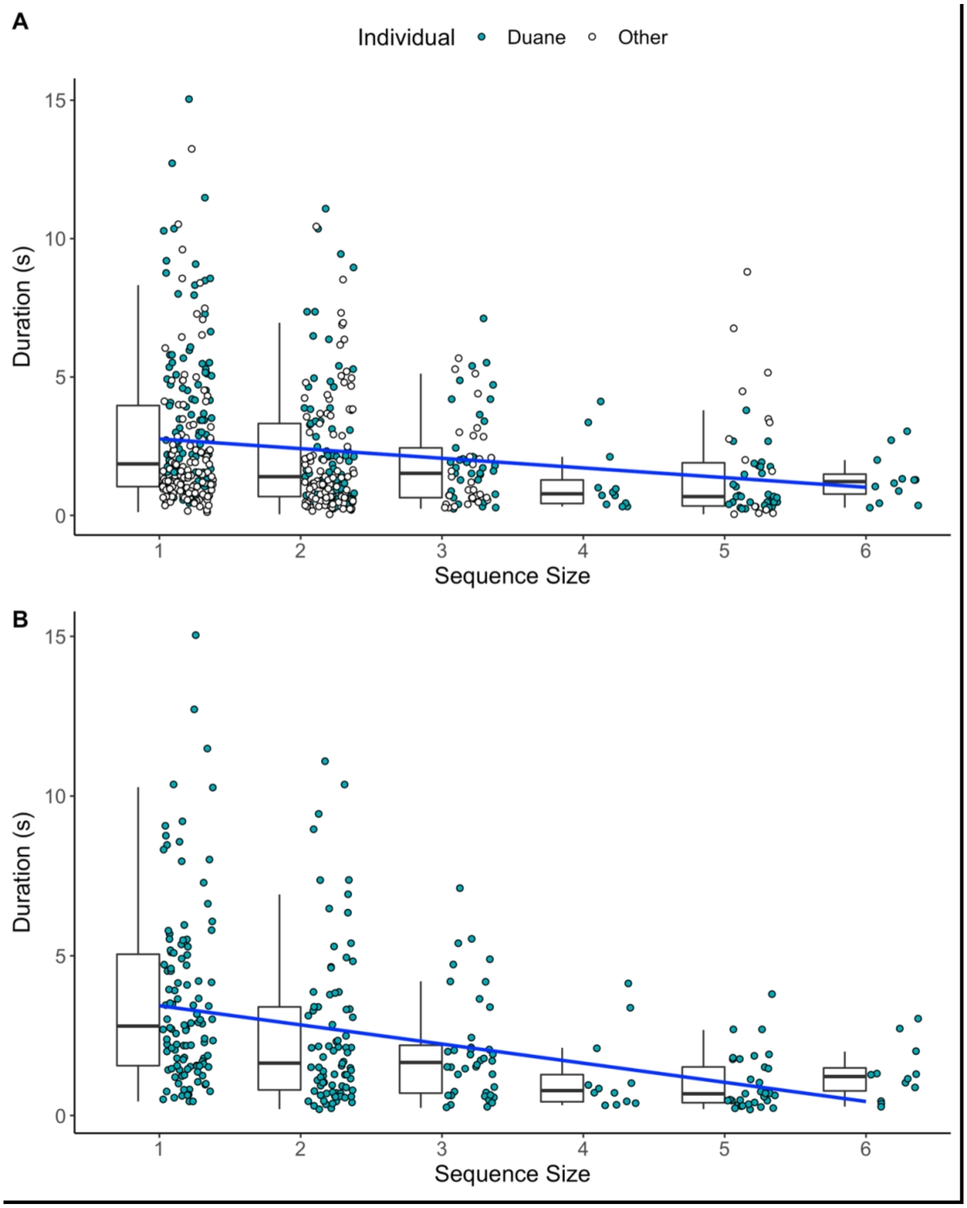


Figure 3. Relationship between sequence size and gesture duration for the full dataset (A) and Duane only data (B). Boxplots represent the median (black bar), the interquartile range – IQR (boxes), and maximum and minimum values excluding outliers (whiskers). Points represent individual gesture tokens, 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.

159x199mm (220 x 220 DPI)

Table 1

Sequence length	***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	

Supporting information 1 – Gesture types definitions

Table S1. Ethogram of the 26 gesture types recorded in the dataset. Definitions are taken from (Hobaiter & Byrne, 2011a) and (Nishida, 2010). Video examples and illustrations of these gestures are available at www.greatapedictionary.com

Gesture	Description	Type
Beckon	Hand is moved in an upwards sweep from the elbow or wrist towards signaller.	Manual
Big loud scratch	Loud exaggerated scratching movement on the signaller's own body.	Manual
Drum	Short hard audible contact of alternate palms against an object, usually tree roots.	Manual
Hit object/ground	Movement of whole arm, with short hard audible contact of closed fist to an object or the ground. Includes gestures performed with one and both arms.	Manual
Hit object/ground with object	As 'hit object/ground' but the signaller holds an object in the hand/hands, which contacts the ground.	Manual
Jump	While bipedal, both feet leave the ground simultaneously, accompanied by horizontal displacement through the air.	Body
Leaf clipping	Strips are torn from a leaf (or leaves) held in the hand using the teeth; produces a conspicuous sound.	Manual
Locomote: bipedal	The signaller walks bipedally while standing up.	Body
Object move	Object is displaced in one direction, contact is maintained throughout movement. Includes gestures performed with one or both hands.	Manual
Object shake	Repeated back and forth movement of an object, usually stem of shrub, branch of tree or woody vine, performed with either one or both hands.	Manual
Present: genitals forwards	Signaller shows genitals to recipient.	Body
Raise arm	Raise arm and/or hand vertically in the air and direct palm to companion.	Manual

Reach: palm	Arm extended to the recipient with the palm exposed. Typically held up or to the side, although very occasionally down. It is the palm or tip of the fingers that is closest to the recipient.	Manual
Reach: wrist	Arm extended to the recipient with the palm sheltered (fingers are curled), and it is either the wrist, or the back of the fingers that is reached out to the recipient.	Manual
Rocking: sitting	Slight or vigorous side to side movements of the body when the signaller is sitting.	Body
Shake arm	Small, repeated shake (adduct or abduct) of horizontally held arm at another. Includes gestures performed with either one or both arms.	Manual
Shake head	Small repeated back and forth motion of the head.	Manual
Stomp 2-feet object/ground	As 'stomp object/ground' but performed with both feet.	Manual
Stomp object/ground	Sole of the foot is lifted vertically and brought into a short hard audible contact with the surface being stood upon (e.g., ground, branch).	Manual
Stomping object/ground	As 'stomp object/ground' but performed repeatedly.	Manual
Swing	Large back and forth movement of the arm held below the shoulder, or of leg from the hip. Includes gestures performed with one and two arms.	Manual
Swing: directed	As 'swing' but the direction of the swing indicates the direction of desired movement, immediately followed by the recipient moving as indicated.	Manual
Swing: with object	As 'swing' but the signaller holds an object in their hand/hands (e.g., branch, leaves, etc).	Manual
Throw object	Object is moved and released so that there is displacement through the air after the moment of release.	Manual
Thrust	Rhythmic back and forth movements of the pelvis.	Body
Wave	Large repeated back and forth movement of the arm raised above the shoulder.	Manual

1

Supporting information 2 - Duration distribution analysis

Before performing the GLMM analysis we analysed the distribution of the gesture duration data by (1) visually inspecting its empirical density and cumulative distribution (Figure S4) and (2) assessing its skewness and kurtosis via the visual inspection of the Cullen and Frey graph (Figure S5). Figure S5 shows the data are skewed towards low values, as almost half of the data lays between 0 and 3 seconds. Further, we fitted three theoretical distributions to the data – namely Weibull, Gamma, and Lognormal – and compared loglikelihood values (Table S2). We then plotted the three distributions and visually inspected the Q-Q, P-P, and histogram density plots (Figure S6). Finally, we compared Weibull, Gamma, and Lognormal distributions against gesture duration data distribution via goodness-of-fit tests and goodness-of-fit information criterion (Table S3 and S4), which helped identify the lognormal distribution as the best fitting one. Therefore, we proceeded with log-transforming the duration variable to best fit model assumptions.

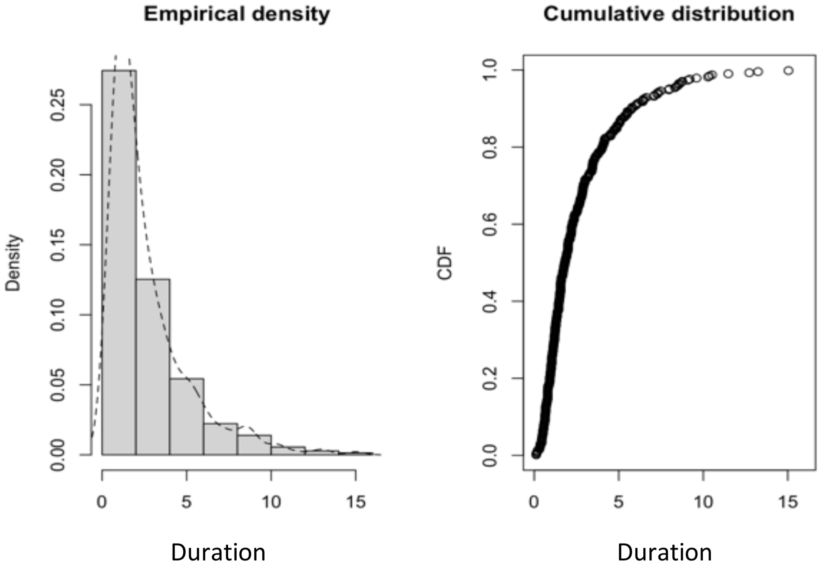


Figure S1. Empirical distribution of gesture duration.

Histogram and empirical cumulative distribution function (CDF) plots representing the distribution of gesture duration. Histogram bars represent sample distribution, dashed line indicates empirical density.

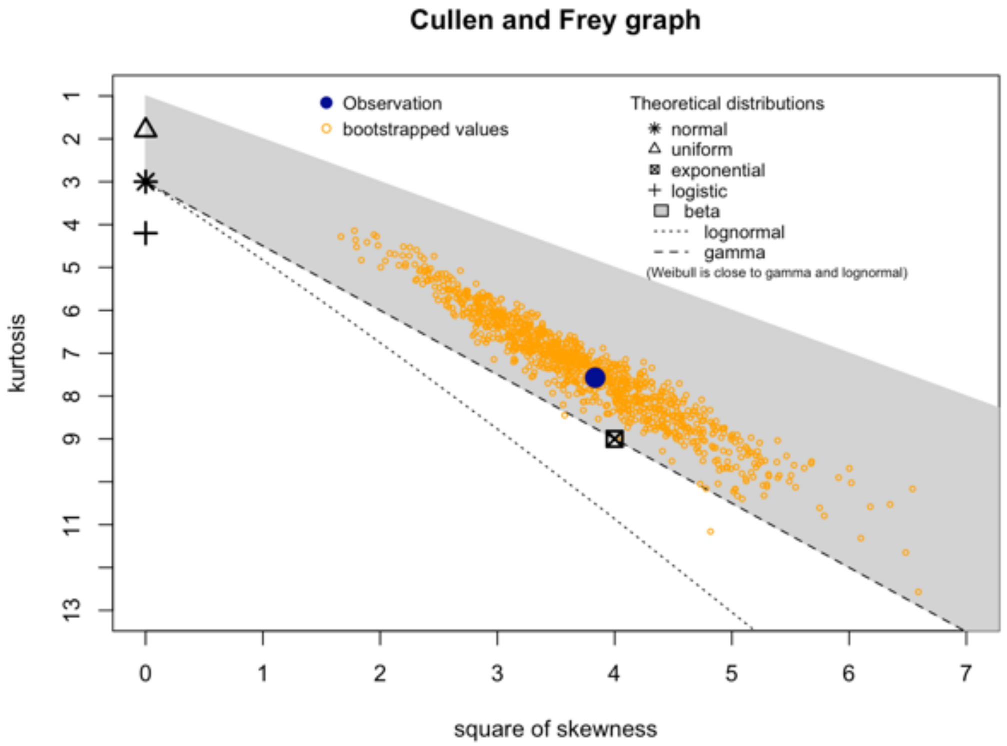


Figure S2. Cullen and Frey graph for gesture duration. The graph depicts the distribution of the skewness and kurtosis of gesture duration data with bootstrapped values, plotted against other theoretical distributions, namely normal, uniform, exponential, logistic, beta, lognormal, and gamma.

Table S2. Estimate and standard error for fitting the parameters of three theoretical distributions to the distribution of the gesture duration data.

Distribution	Parameters	Estimate	Std Error	Loglikelihood
Weibull	Shape	1.229711	0.04786805	-695.5846
	Scale	2.848181	0.12958597	
Gamma	Shape	1.5695397	0.10700744	-688.7789
	Rate	0.5933806	0.04755517	
Lognormal	meanlog	0.6214851	0.04530977	-677.7389
	sdlog	0.8584975	0.03203865	

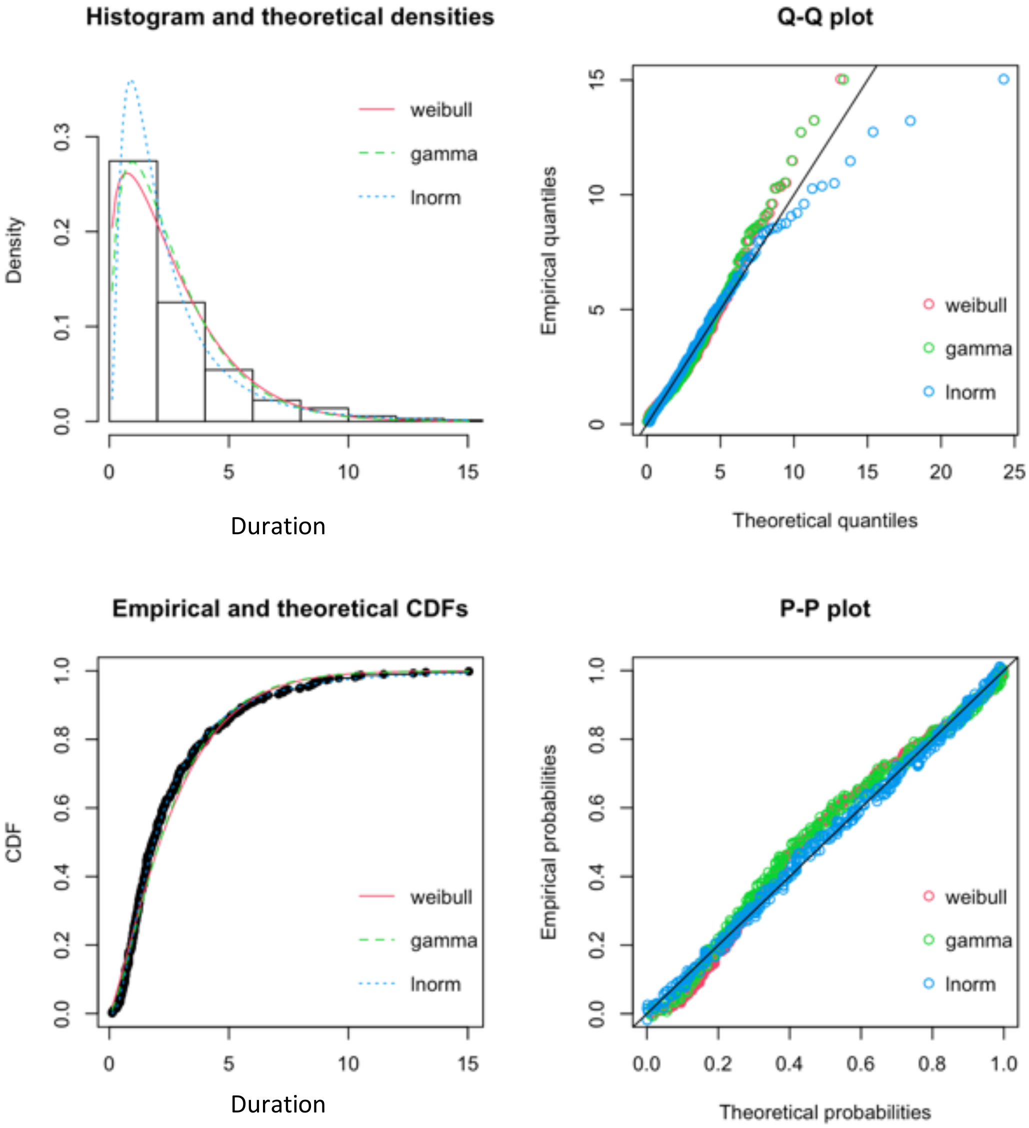


Figure S3. Histogram and theoretical densities, Q-Q and P-P plots depicting the gesture duration data distribution against the fitted Weibull, Gamma, and Lognormal distributions.

Histogram represents the distribution of duration data while the red, dashed green, and dashed blue lines indicate the theoretical Weibull, Gamma, and Lognormal distributions, respectively.

Table S3. Goodness-of-fit statistics compared across fitted distributions to the gesture duration data.

Goodness-of-fit statistics	Weibull	Gamma	Lognormal

Kolmogorov-Smirnov statistic	0. 07621865	0.08155393	0. 03242513
Cramer-von Mises statistic	0.62478933	0.55850079	0.03735280
Anderson-Darling statistic	3.88457421	3.08240994	0.30076490

Table S4. Goodness-of-fit information criteria compared across fitted distributions.

Goodness-of-fit criteria	**Weibull**	**Gamma**	**Lognormal**
Akaike's Information Criterion	1395.169	1381.558	1359.478
Bayesian Information Criterion	1402.936	1389.324	1367.244

Supporting information 3 - Model results

Table S5. Summary of the Bayesian mixed model analysis results for the Zipf-model which included all the data (*N*=560).

Fixed effects	*b*	*SD*	95% CrI	*Bulk ESS*	*Tail ESS*	$\hat{R}$
Intercept	0.02	0.36	[-0.67; 0.72]	1257	1528	1.00
P	0.90	1.26	[-1.25; 3.81]	1298	1173	1.00
Category						
Whole body	*Reference*					
Manual	-0.24	0.39	[-1.04; 0.49]	966	1762	1.00
Random effects						
Gesture, *N*=26	0.80	0.14	[0.57; 1.13]	809	1275	1.00
Sequence ID, *N*=377	0.10	0.07	[0.01; 0.25]	459	960	1.00
Signaller ID, *N*=16	0.20	0.06	[0.11; 0.34]	1749	1936	1.00

Table S6. Summary of the Bayesian mixed model analysis results for the Zipf-model which included only Duane's data (*N*=290).

Intercept	-0.48	0.50	[-1.46; 0.55]	1867	1715	1.00
P	1.79	1.15	[-0.39; 4.00]	1029	1547	1.00
Category						
Whole body	*Reference*					
Manual	0.31	0.50	[-0.72; 1.30]	1663	1672	1.00
Date						
03/02/2008	*Reference*					
05/01/2008	-0.20	0.11	[-0.43; 0.02]	3183	2572	1.00
20/01/2008	0.12	0.11	[-0.09; 0.35]	2765	2161	1.00
Random effects						
Gesture, *N*=15	0.49	0.17	[0.23; 0.89]	867	1541	1.00
Sequence ID, *N*=181	0.11	0.07	[0.01; 0.26]	775	1328	1.00

Table S7. Summary of the Bayesian mixed model analysis results for the Zipf-model which included data from all individuals but Duane (*N*=270).

Fixed effects	*b*	*SD*	95% CrI	*Bulk ESS*	*Tail ESS*	$\hat{R}$
Intercept	0.27	0.46	[-0.61; 1.21]	1442	1499	1.00
P	0.00	0.00	[-0.00; 0.01]	2428	1993	1.00
Category						
Whole body	*Reference*					
Manual	-0.36	0.50	[-1.38; 63]	1565	1769	1.00
Random effects						
Gesture, *N*=18	0.87	0.19	[0.58; 1.32]	1194	1657	1.00
Sequence ID, *N*=196	0.20	0.10	[0.01; 0.38]	468	1053	1.01

Table S8. Summary of the Bayesian mixed model analysis results for the Menzerath-model which included all the data (*N*=530).

Fixed effects	*b*	*SD*	95% CrI	*Bulk ESS*	*Tail ESS*	$\hat{R}$
Intercept	0.69	0.14	[0.41; 0.97]	1316	1809	1.00
Sequence Size	**-0.18**	**0.04**	**[-0.26; -0.11]**	**2374**	**1915**	**1.00**
PWB	-0.31	0.20	[-0.71; 0.08]	2795	2529	1.00
Random effects						
Signaller ID, N=16	0.36	0.11	[0.18; 0.62]	938	1501	1.00
Sequence ID, N=359	0.31	0.10	[0.07; 0.47]	356	539	1.01

Table S9. Summary of the Bayesian mixed model analysis results for the Menzerath-model which included only Duane's data (*N*=273).

Fixed effects	***b***	***SD***	**95% CrI**	***Bulk ESS***	***Tail ESS***	$\hat{R}$
Intercept	0.92	0.13	[0.67; 1.17]	4365	2026	1.00
Sequence Size	**-0.23**	**0.04**	**[-0.31; -0.15]**	**4074**	**2145**	**1.00**
PWB	-0.69	0.58	[-1.89; 0.38]	4900	2168	1.00
Date						
03/02/2008	*Reference*					
05/01/2008	0.06	0.14	[-0.22; 0.34]	4821	2130	1.00
20/01/2008	0.48	0.13	[0.22; 0.75]	5325	2610	1.00
Random effects						
Sequence ID, N=181	0.11	0.08	[0.00; 0.30]	995	1009	1.00

Table S10. Summary of the Bayesian mixed model analysis results for the Menzerath-model which included data from all individuals but Duane (*N*=257).

Fixed effects	***b***	***SD***	**95% CrI**	***Bulk ESS***	***Tail ESS***	$\hat{R}$
Intercept	0.33	0.18	[-0.04; 0.70]	2167	2241	1.00
Sequence Size	0.01	0.08	[-0.14; 0.16]	2866	2438	1.00
PWB	-0.24	0.22	[-0.69; 18]	3474	2226	1.00
Random effects						
Sequence ID, *N*=187	0.42	0.11	[0.16; 0.62]	466	423	1.01
Signaller ID, *N*=15	0.37	0.14	[0.14; 0.70]	842	902	1.00

Abbreviations: *b*= Estimated mean of the posterior distribution; *SD*= Standard deviation of the posterior distribution; *CrI*= Two-sided 95% Credible intervals based on quantiles; *Bulk ESS*= the effective sample size for rank normalized values using split chains; *Tail ESS*= the minimum of the effective sample sizes for 5% and 95% quantiles; *R*ˆ=R hat value, provides information about the convergence of the Bayesian model algorithm.

Supporting information 4

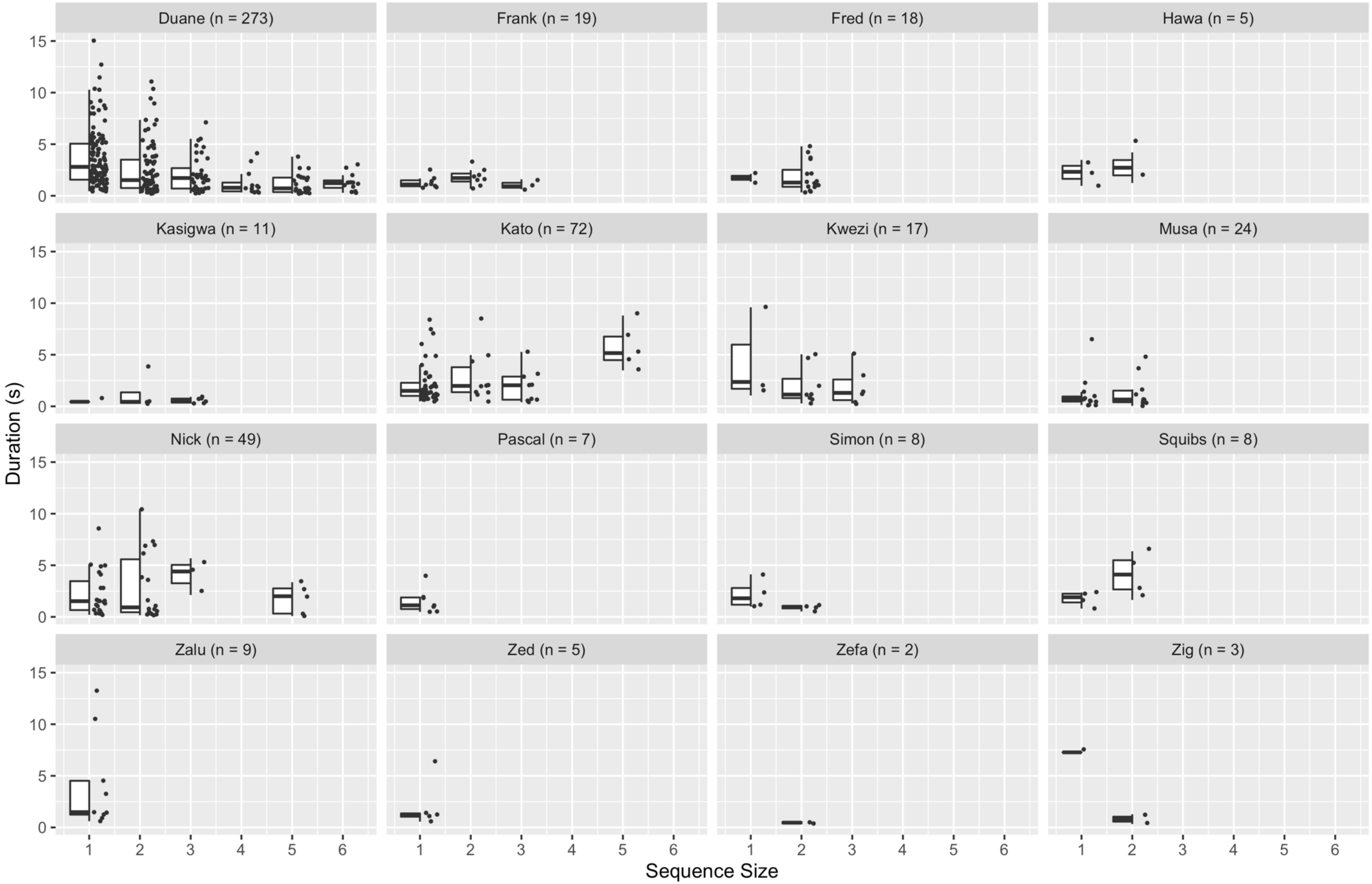


Figure S4. Distribution of gesture durations based on sequence size for each of the 16 individuals in the dataset. Points represent individual gesture tokens. Boxplots show median (black central bar), interquartile range (boxes), maximum and minimum values exploding outliers (whiskers). n indicates sample size for each individual.

Supporting information 5 – Visual inspection of sequence structure

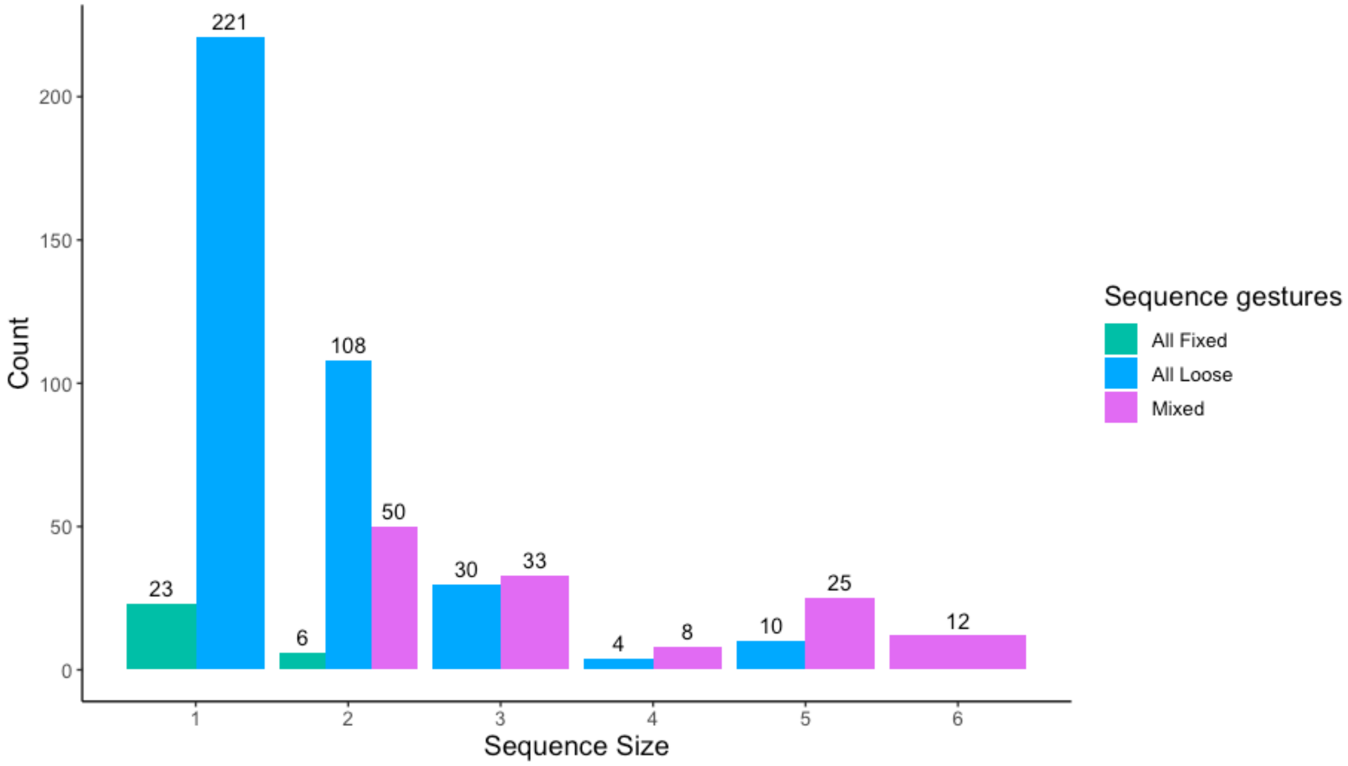


Figure S5. Bar chart showing the frequency distribution of the three different types of sequences depending on their sequence size.

Green sequences comprise only *fixed* duration gestures, blue sequences only *loose* duration gestures. Pink bars represent sequences formed by a mix of *loose* and *fixed* duration gestures. Numbers above bars indicate the frequency of each sequence type per sequence size.

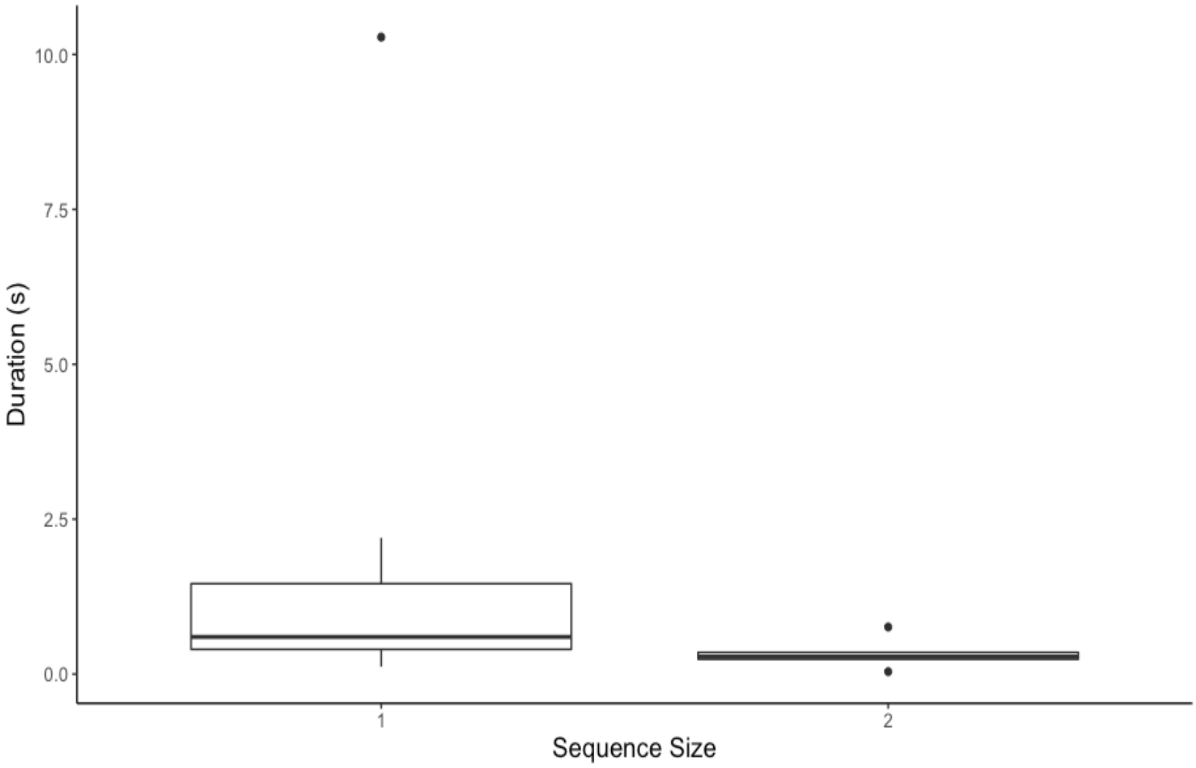


Figure S6. Boxplots of the duration of gestures with constrained duration (i.e., *fixed* duration gestures) in sequences formed of *fixed* duration gestures solely.

Boxplots show median (black central bar), interquartile range (boxes), maximum and minimum values exploding outliers (whiskers) and outliers (circles).

Supporting information 6 – Alternative analyses

Correlation and compression

Methods

Compression predicts that mean duration should be smaller than expected by chance (Ferrer-i-Cancho et al., 2013). Similarly, optimal compression predicts linguistic laws as a correlation in a specific direction, *i.e.,* the correlation cannot be positive (Ferrer-i-Cancho et al., 2013, 2020). Accordingly, we employed one-tailed tests of compression throughout, but we also report the outcome of two-tailed equivalents for comparison with previous findings (Heesen et al., 2019).

We conducted one-tailed Spearman rank correlation tests to analyse the relationship between the frequency within the sample of a gesture type (*frequency*) and its mean duration (*mean gesture type duration*), calculated by dividing the total sum of all durations of the same gesture type (Sum), by *frequency (i.e., duration=Sum/frequency)* (Semple et al., 2013). A similar procedure was used to test for a correlation between the mean gesture duration within a given sequence (*sequence*) and the number of gesture tokens in the same sequence (*n*). Mean gesture duration was calculated by dividing the total duration of a gestural sequence (*Total*) – *i.e.*, the sum of all durations of the gesture tokens in the sequence excluding pauses between gestures – by the number of gesture tokens within that sequence *n* (*i.e., mean gesture duration within sequence=sum of durations of gestures within the sequence/number of gestures within the sequence*). A negative correlation between *mean gesture type duration* and *frequency* coherent with Zipf's law of brevity, and a negative correlation between *t* and *n* conforming to Menzerath's law could both be unavoidable artefacts given the relationship between *d* and *f,* and between *t* and *n* – as defining *d* involves *f,* and defining *t* involves *n* – which could lead to *d* = 1/*f* and *t*=1/*n* (Ferrer-i-Cancho et al., 2014). Such artefacts can be excluded by establishing that *D* and *f,* and *T* and *n* are significantly positively correlated (Ferrer-i-Cancho et al., 2014; Semple et al., 2013), which we tested using two Spearman rank correlation tests. Current findings suggest that the expressions of linguistic laws are not 'universal', and there may be more variation than previously recognised (Semple et al., 2022). For example: earlier research demonstrated Zipf's law of brevity can be present in parts of a repertoire, when it appears to be absent in the whole repertoire (Ferrer-i-Cancho & Hernández-Fernández, 2013; Heesen et al., 2019). As a result, we also tested for Zipf's law of brevity in specific subsets of the repertoire, namely manual versus whole-body gesture types as these had been found to differ in previous work (Heesen et al., 2019). Moreover, a specific check of Zipf's law of brevity in manual gestures aids in comparison with studies of human communication that only consider manual signals (for example in signed languages and fingerspelling).

Compression test

Is the mean duration of chimpanzee sexual solicitation gesture types significantly small?

Following earlier work on chimpanzee play gestures (Heesen et al., 2019), we first calculated mean duration of all gesture types *L* via the equation:

$$L = \sum_{i=1}^{n} p_i e_i \qquad (1)$$

where *n* is the number of elements within the repertoire, p_i is the normalized probability of the i^{th} element – calculated by dividing the frequency of the i^{th} gesture by the total frequency of all gestures – and e_i is the magnitude of the i^{th} element (*i.e.*, its average duration *d*).

To test for compression and whether Zipf's law holds in chimpanzee sexual solicitation gestural communication, we used a permutation test assessing whether *L* was significantly small (Sokal & Rohlf, 1995). Following (Heesen et al., 2019) we created "*a control distribution of L (L') defined by a permutation function π (i)"* and calculated the left *p*-value by dividing the number of permutations where *L'≤L* by the number of total permutations, here 10^5. *L* was also calculated and tested for each subset created.

$$L' = \sum_{i=1}^{n} p_i e_{\pi(i)} \qquad (2)$$

Is the expected total sum of the duration of gestures of each sequence significantly small?

As explained by (Heesen et al., 2019), the total duration of a collection of sequences can be quantified as

$$M = \sum_{i=1}^{N} T_i \qquad (3)$$

where T_i is the total duration of the ith sequence and *N* is the number of sequences.

In turn,

$$T = \sum_{j=1}^{n_i} t_{ij} \qquad (4)$$

where t_{ij} is the duration of the jth element of the ith sequence and n_i is the size of the ith sequence. Given that the mean duration of gestures from the ith sequence can be expressed as $\left(t_{ij}\right)_i = \frac{T_i}{n_i}$, M can be defined as

$$M = \sum_{i=1}^{N} n_i \left(t_{ij}\right)_i \tag{5}$$

M was calculated through this equation and was tested to assess whether it is significantly small. We performed a similar permutation test to that conducted to test for the significance of *L*, to check whether *M* was significantly small as compared to the values generated by random permutation of the data (Zipf, 1936). In such case, n_i has the role of p_i and $\left(t_{ij}\right)_i$ has the role of e_i in the test, with n_i and $\left(t_{ij}\right)_i$ remaining constant during the test.

The permutation test produces a left *p*-value to check if *L* (or *M*) is significantly small and a right *p*-value to check if *L* (or *M*) is significantly large compared to the distribution of the values created by a permutation of the data (Heesen et al., 2019). The total number of permutations carried out was $R=10^5$.

Results of one-tailed analyses

Zipf's law of brevity

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

We did not find a pattern in agreement with Zipf's law of brevity; there was no evidence for a significant negative correlation between mean gesture type duration (*d*) and frequency of use (*f*) (Spearman correlation: $r_s=0.30$, $n=26$, $p=0.066$), in agreement with the Bayesian model analysis. Consistent with this result, the compression test revealed that the expected mean code length of gesture types *L* had a magnitude of 2.39s and was not significantly small ($p_{left}=0.951$). Rather, *L* was significantly big ($p_{right}=0.05$, Figure S7).

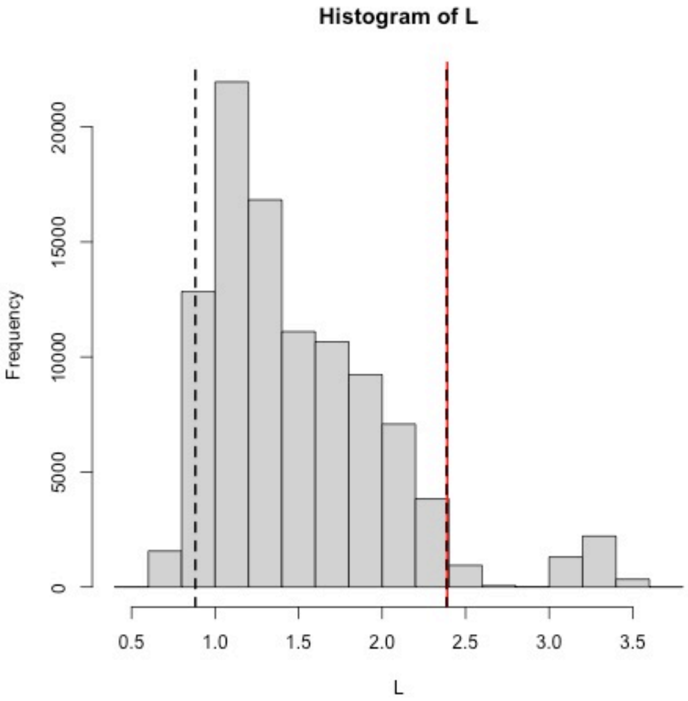


Figure S7. Histogram showing the distribution of the permuted *L* values. Observed *L* value is highlighted with the red continuous line. Black dashed line indicates the lower and upper 5% of the permuted data.

Subset analysis: whole-body and manual gesture types.

We found no evidence for a negative correlation between *d* and *f* when separating whole-body gestures from manual gestures (Spearman's rank correlation: whole-body, r_s=-0.3, *n*=5, *p*=0.342; manual, r_s =0.42, *n*=21, p_{left}=0.969). Rather, manual gestures showed a significant positive correlation (r_s =0.42, *n*=21, p_{right}=0.031). Compression tests revealed that for whole-body gestures, *L*=0.13s and was neither significantly big or small (p_{left}=0.174, p_{right}=0.817), and for manual gestures, *L*=2.26s and, if anything, tended towards being significantly big (p_{right}=0.058) rather than small (p_{left}=0.942; Figure S8).

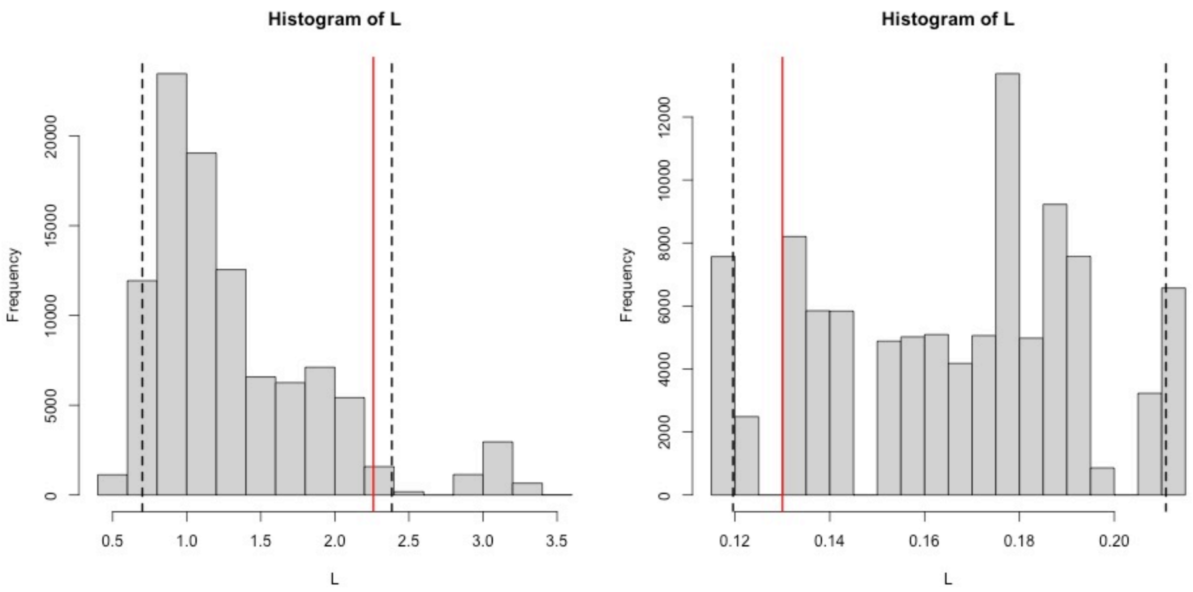


Figure S8. Histograms showing the distribution of the permuted *L* values for manual gestures (left) and whole-body gestures (right). Observed *L* value is highlighted with the red continuous line. Black dashed lines indicate the upper and lower 5% of the permuted data.

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

We tested Menzerath's law in 359 sequences, composed of 530 gesture tokens; there was no evidence for a negative relationship between mean constituent duration and sequence size (Spearman's rank correlation: r_s=-0.08 *n*=359, *p*=0.076). However, the compression test revealed that the total sum of the duration of each sequence *M* had a value of 1300.67 and was significantly small (*n*=359, *p*=0.003; Figure S9).

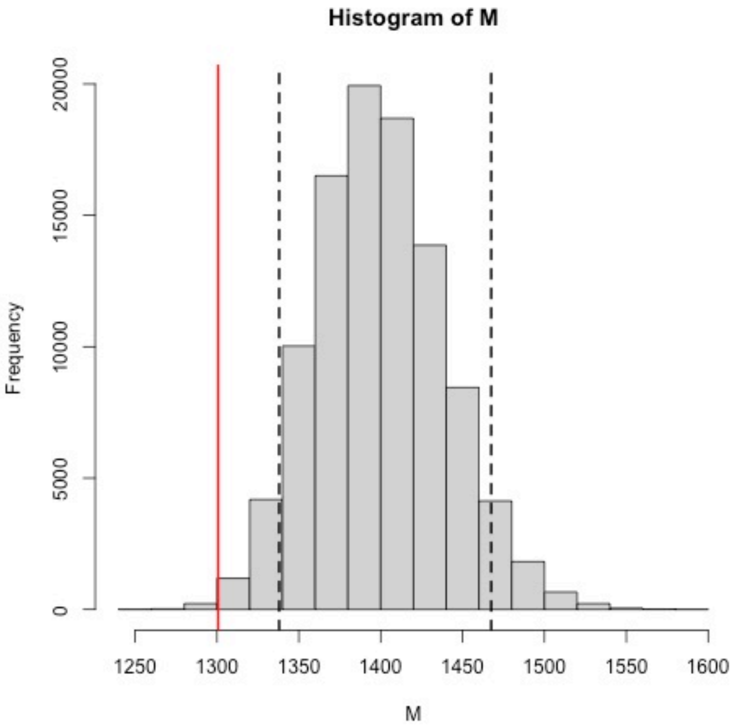


Figure S9. Histogram showing the distribution of the permuted *M* values. Observed *M* value is highlighted with the red continuous line. Black dashed line indicates lower 5% of the permuted data.

Discussion

The results from the correlation analysis must be taken with caution as this analysis does not control for individual variation, gesture type, and sequence in which the gesture is performed. The Bayesian model that included these factors and which tested for Zipf's law of brevity (Zipf-model) was similar to the respective null-model, suggesting that frequency of gesture type within the dataset and category of gesture type did not predict gesture duration.

The contrast between the correlation analysis and the compression analysis for Menzerath's law highlight how individual variation may show an apparent absence of pattern in the correlation analysis but a strong effect in the Bayesian model analysis, where it is controlled for.

Results of two-tailed analyses

Zipf's law of brevity

We did not find a pattern corresponding to Zipf's law of brevity, with no correlation between mean gesture type duration (d) and frequency of use (f) (Spearman correlation: r_s=0.30, n=26, p=0.131). When analysing only manual gestures, f and d tended to be significantly positively correlated (Spearman correlation: r_s=0.42, n=21, p=0.061). Conversely, we did not find any correlation between f and d in whole body gestures (Spearman correlation: r_s=-0.3, n=5, p=0.683).

Menzerath's law

We failed to find a pattern between sequence size n and mean constituent duration t of the same sequence that followed Menzerath's law (Spearman correlation: r_s=-0.08, n=376, p=0.142). When analysing sequences comprising only whole-body size and average gesture duration showed a significant positive correlation (Spearman correlation: r_s=0.59, n=20 p=0.005). Sequence size and average gesture duration did not correlate in sequences composed of only manual gestures (Spearman correlation: r_s=-0.06, n=315 p=0.324), or in those formed by both manual and body gestures (Spearman correlation: r_s=0.09, n=24 p=0.673).

Bayesian analysis on the medians of the 26 gesture types

We ran additional brms analysis computing the median duration per each gesture type across the whole dataset. Median duration of each gesture type was assigned as response variable, category of gesture as fixed factor as well as frequency of that gesture type as a predictor. We ran a model with 2000 iterations and 3 chains. The full model was no different from the null model which excluded the frequency of gesture type as a predictor (LOO difference: -0.3 ± 0.5; Table S11). We ran a similar analysis on the gestures performed by the one individual Duane, with similar results (full-null model comparison, LOO difference: -0.6± 1.2; Table S12). Please note that that individual identity is not controlled for in these analyses and they should be interpretted with caution.

Table S11. Summary of the Bayesian mixed model analysis results for the Zipf-model which included only a median value per gesture type as response variable, the frequency of gesture type as predictor and category of gesture type as control (*N*=26).

Fixed effects	***b***	***SD***	**95% CrI**	***Bulk ESS***	***Tail ESS***	***$\hat{R}$***
Intercept	0.10	0.33	[-0.55; 0.77]	2267	1782	1.00
F	0.00	0.00	[-0.00; 0.01]	2861	1404	1.00
Category						
Whole body	*Reference*					
Manual	-0.24	0.36	[-0.96; 0.48]	1995	1776	1.00

Table S12. Summary of the Bayesian mixed model analysis results for the Zipf-model which included only a median value per gesture type as response variable, the frequency of gesture type as predictor and category of gesture type as control, considering only the gestures performed by Duane (*N*=15).

Fixed effects	***b***	***SD***	**95% CrI**	***Bulk ESS***	***Tail ESS***	***$\hat{R}$***
Intercept	-0.55	0.57	[-1.71; 0.60]	2254	1851	1.00
F	0.01	0.00	[-0.00; 0.02]	2566	1771	1.00
Category						
Whole body	*Reference*					
Manual	0.26	0.57	[-0.89; 1.48]	2334	1714	1.00