Supplemental information

Chimpanzees show the capacity to communicate

about concomitant daily life events

Tatiana Bortolato, Angela D. Friederici, Cédric Girard-Buttoz, Roman M. Wittig, and Catherine Crockford
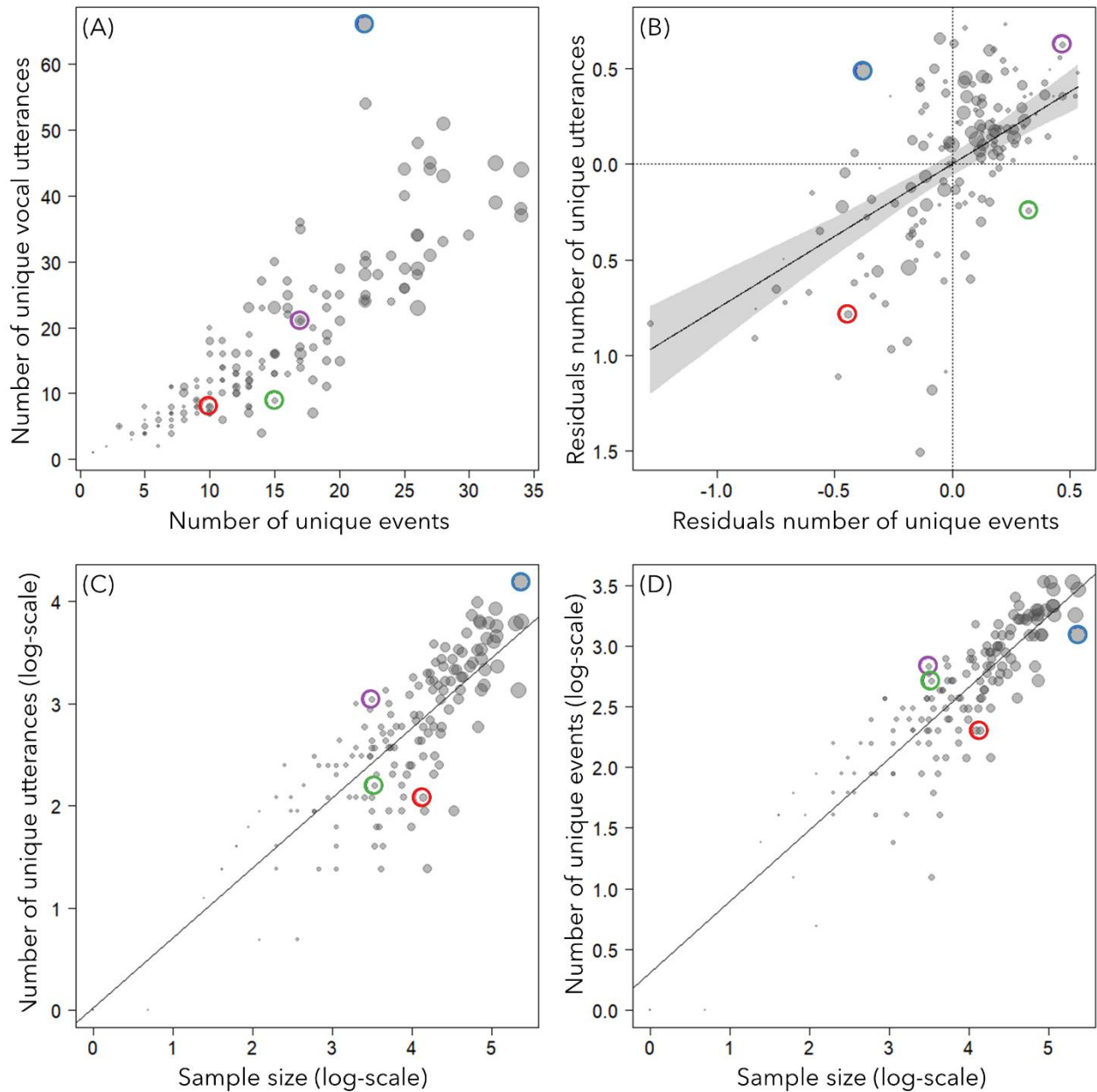
Figure S1: Illustration of the rationale of Model 2a, related to STAR Methods. (A) Relation between the number of unique vocal utterances and the number of unique events (including single and concomitant events). (B) Effect of the residuals extracted from the model from plot (D) on the residuals extracted from the model in plot (C). (C) Effect of sample size using log-scale on the number of unique vocal utterances also on log-scale. (D) Effect of sample size using log scale on the number of unique events also using log-scale. Each dot represents one individual per season, and their area is proportionate to the total number of utterances per chimpanzee and field season. Coloured circles show specific individuals across the figures for a better interpretation of the plots. Black lines depict model lines and the polygon its 95% confidence limits. Grey dashed lines mark the zero of both residuals. Data points above the model line in (c) (e.g., purple and blue dots) will be above the grey dashed line on the y-axis of the plot (B), and data points below the model line in (C) (e.g., green and red dots) will appear below the grey dashed line on the y-axis in the plot (B). The same rationale works for dots above or below the model line in plot (D) regarding the x-axis of plot (B).
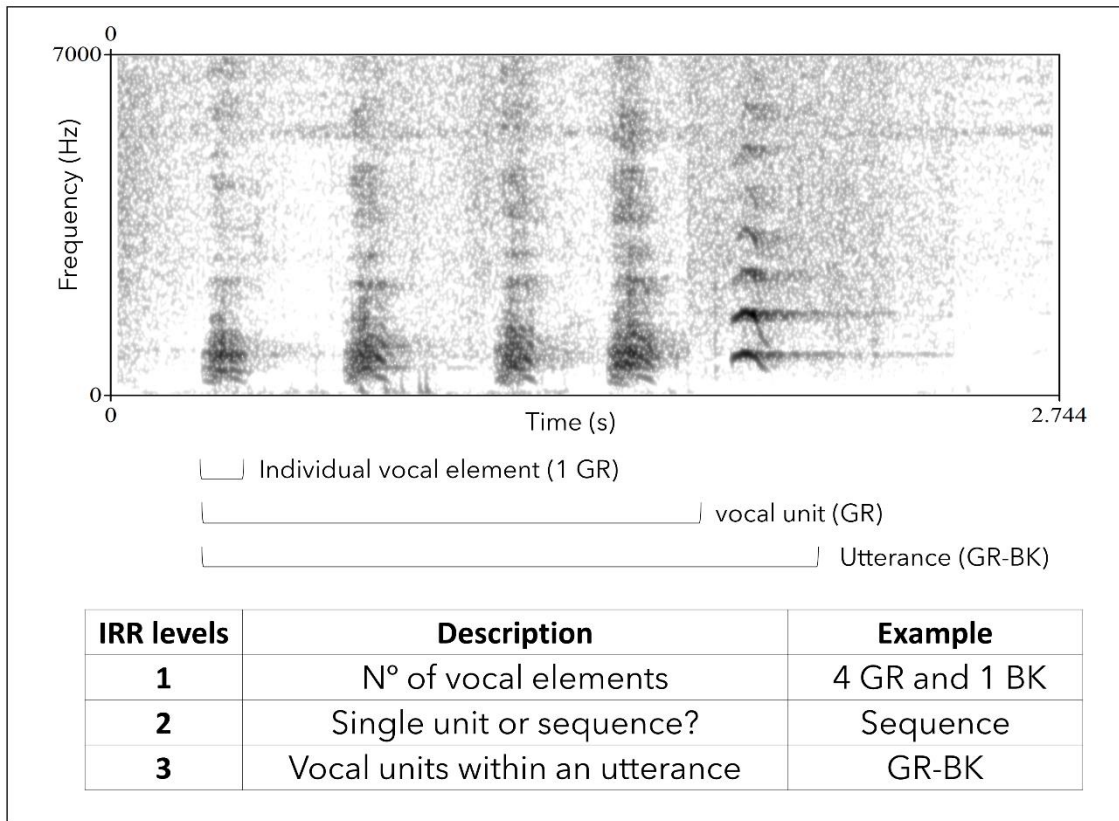
IRR levels	Description	Example
1	N° of vocal elements	4 GR and 1 BK
2	Single unit or sequence?	Sequence
3	Vocal units within an utterance	GR-BK

Figure S2: Illustration of the inter-rater reliability (IRR) test on vocal coding, related to STAR Methods. The spectrogram depicts an utterance formed by two vocal units (or 'call types'), grunts and a bark. The utterance has a total of 5 vocal elements, the first 4 elements belong to the first vocal unit (grunts) and the last element to the second vocal unit (bark). For the IRR, each coder assessed the number of elements of the different vocal units within the utterance (level 1), whether the utterance was a single vocal unit of a combined sequence (level 2), and the code for the utterance, i.e., the different vocal units within the utterance (level 3).
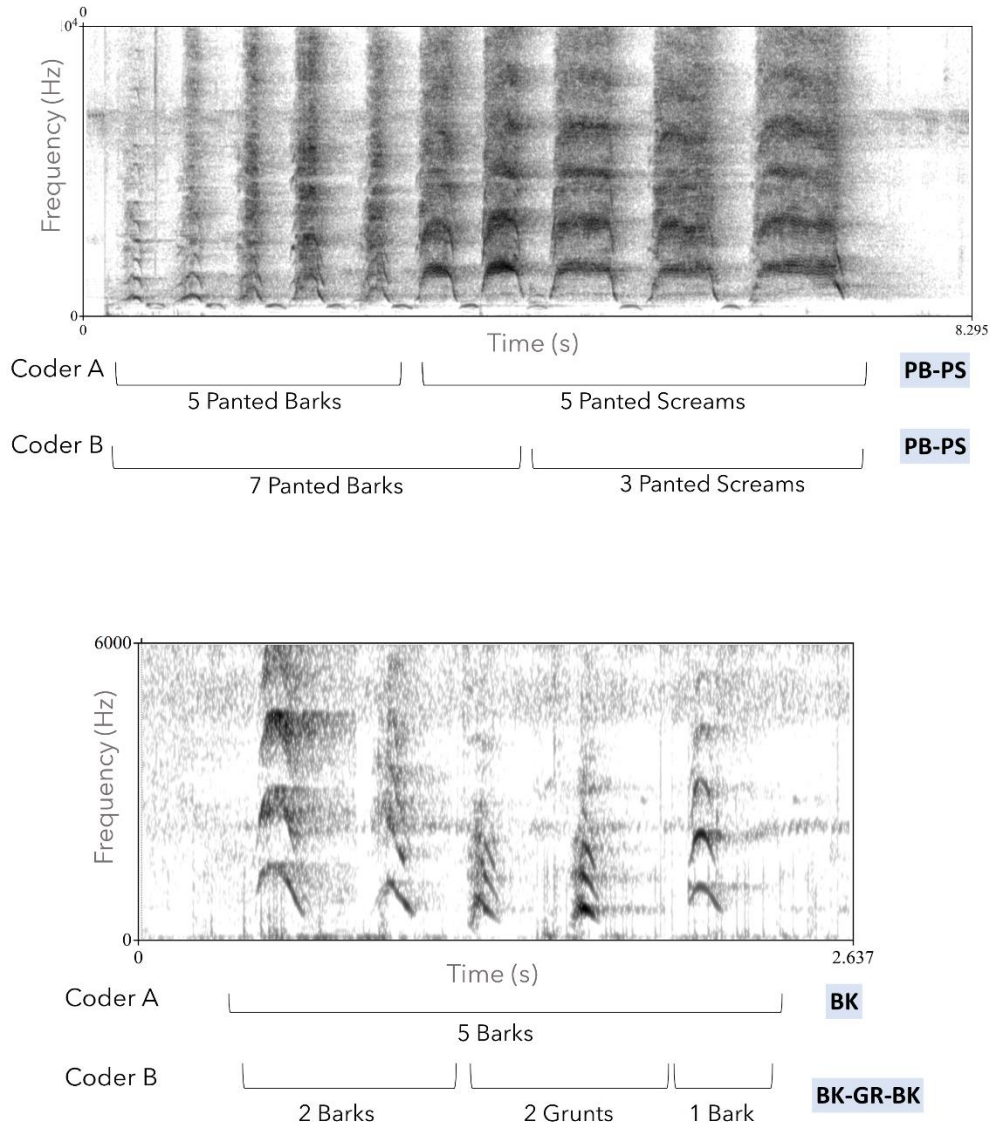
Figure S3: Illustration of the rationale for testing IRR at three levels, related to STAR Methods.
Examples where different assessments of elements can lead to the same or different utterances. On the top, there is a disagreement on the coding of elements, but the utterance is the same between coders (marked in bold and blue). On the bottom, there is an agreement on 3 elements (3 barks) and a disagreement on the other two elements. Moreover, in this example, the utterance is different in both their component units and whether it is a single unit or a sequence. In the cases with an agreement between all the elements, we will also have an agreement between the utterances.

	HO	GR	BK	SC	WH	NV	PN	PH	PG	PB	PS	PW
HO	41	0	0	0	0	1	0	1	0	0	0	0
GR	6	222	0	0	0	0	6	0	8	0	0	0
BK	0	0	50	5	0	0	0	0	0	0	0	0
SC	0	0	1	142	8	0	0	0	0	0	2	0
WH	3	0	0	0	49	0	0	0	0	0	0	0
NV	0	0	0	0	0	30	0	0	0	0	0	0
PN	0	6	0	0	0	0	225	17	3	0	0	0
PH	0	0	0	0	0	0	0	265	0	2	0	0
PG	1	4	0	0	0	0	14	11	135	0	0	0
PB	0	0	2	0	0	0	0	1	0	58	0	0
PS	0	0	0	0	0	0	0	0	0	0	40	0
PW	0	0	0	0	0	0	0	0	0	0	0	2

Table S1: Matrix with the number of elements of each call type, used to measure Cohen's Kappa score between coders, related to STAR Methods. Agreement per element is shown on the diagonal of the matrix (highlighted in yellow). Disagreements are marked outside the diagonal of the matrix.

Table S3: The effect of concomitant events and group on the production of vocal sequences for *adult* individuals (>10y), Model 1b, N=5399, related to STAR Methods.

Term	Estimate	SE	Lower 95% CI	Upper 95% CI	$X2$*	P
Intercept	-0.831	0.079	-0.971	-0.670		
Concomitant event (yes)	1.090	0.089	0.914	1.264	71.945	**<0.001**
Group (North)	-0.257	0.125	-0.465	-0.033	6.728	**0.035**
Group (South)	-0.289	0.116	-0.515	-0.070		

In bold statistically significant results ($P \leq 0.05$). CI, confidence interval. The coded level for each categorical predictor is indicated in parentheses. * Degrees of freedom are 1.

Table S4: The effect of the number of unique *single* events and group on the number of unique utterances emitted per individual, Model 2b, N=152, related to STAR Methods.

Term	Estimate	SE	Lower 95% CI	Upper 95% CI	$X2$*	P
Intercept	0.093	0.056	-0.014	0.196		
Residuals number unique events	0.582	0.101	0.387	0.782	33.029	**<0.001**
Group (North)	-0.094	0.096	-0.271	0.082	2.455	0.091

| Group (South) | -0.173 | 0.078 | -0.305 | -0.025 | | |

In bold statistically significant results ($P \le 0.05$). CI, confidence interval. The coded level for each categorical predictor is indicated in parentheses. * Degrees of freedom are 1.

Table S5: The effect of the number of unique events and group on the number of unique utterances emitted per *adult* individual, Model 2c, N=73, related to STAR Methods.

Term	Estimate	SE	Lower 95% CI	Upper 95% CI	*X2**	*P*
Intercept	0.048	0.038	-0.027	0.123		
Residuals number unique events	0.386	0.107	0.173	0.603	13.009	**<0.001**
Group (North)	-0.061	0.075	-0.197	0.086	1.681	0.195
Group (South)	-0.11	0.060	-0.226	0.016		

In bold statistically significant results ($P \le 0.05$). CI, confidence interval. The coded level for each categorical predictor is indicated in parentheses. * Degrees of freedom are 1.