

Efficient communication in written and performed music

Laurent Bonnasse-Gahot

Supplementary Information

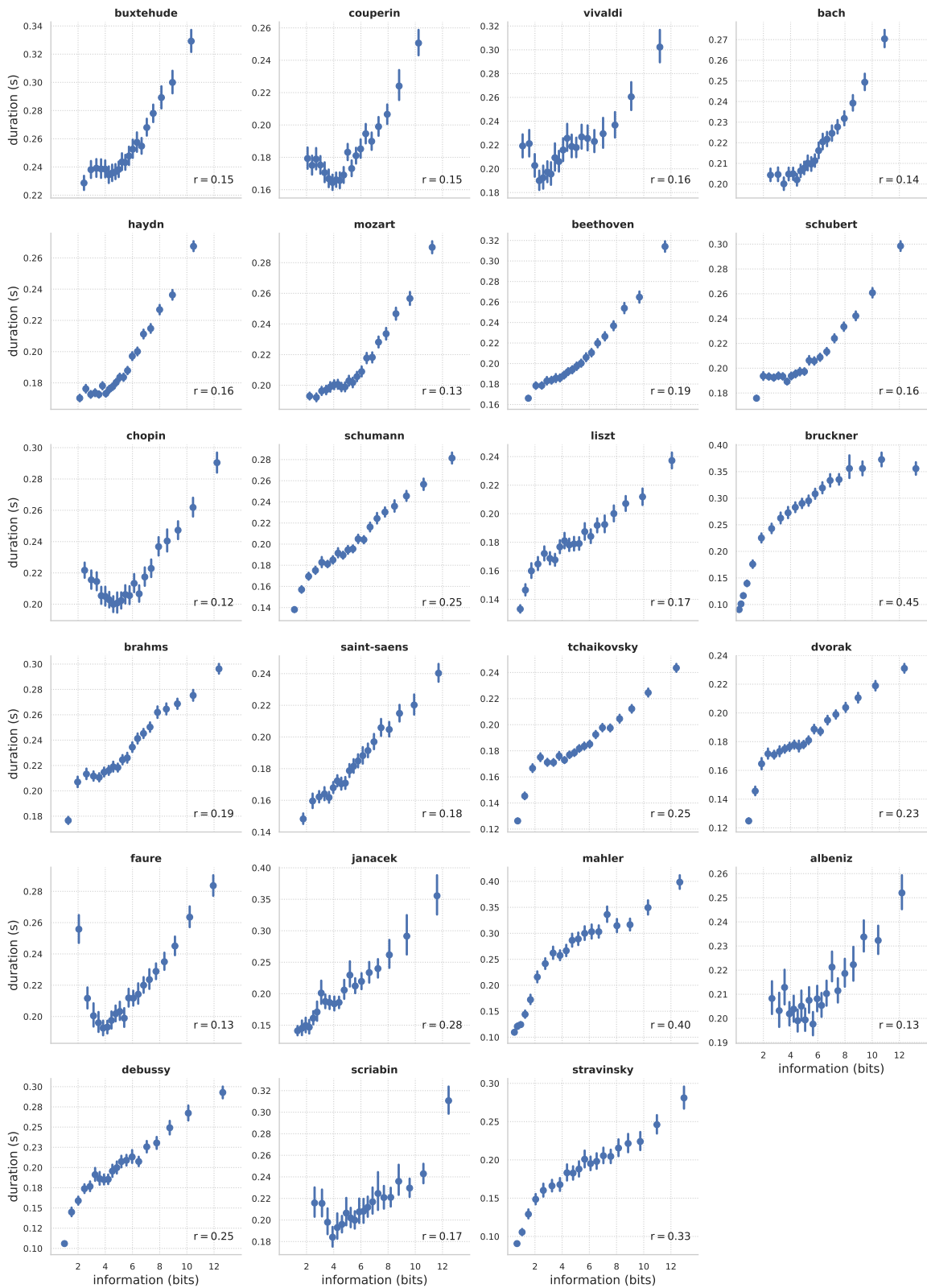


Figure S1: Absolute duration, in seconds, as a function of information content, for each composer in the sequenced corpus. Corresponding Spearman correlation is indicated in each subplot. Composers are ordered according to their date of birth. Each bin represents 5% of the data. Error bars represent 95% confidence intervals, estimated by bootstrap (a given bar might not be visible if below the size of its corresponding marker).

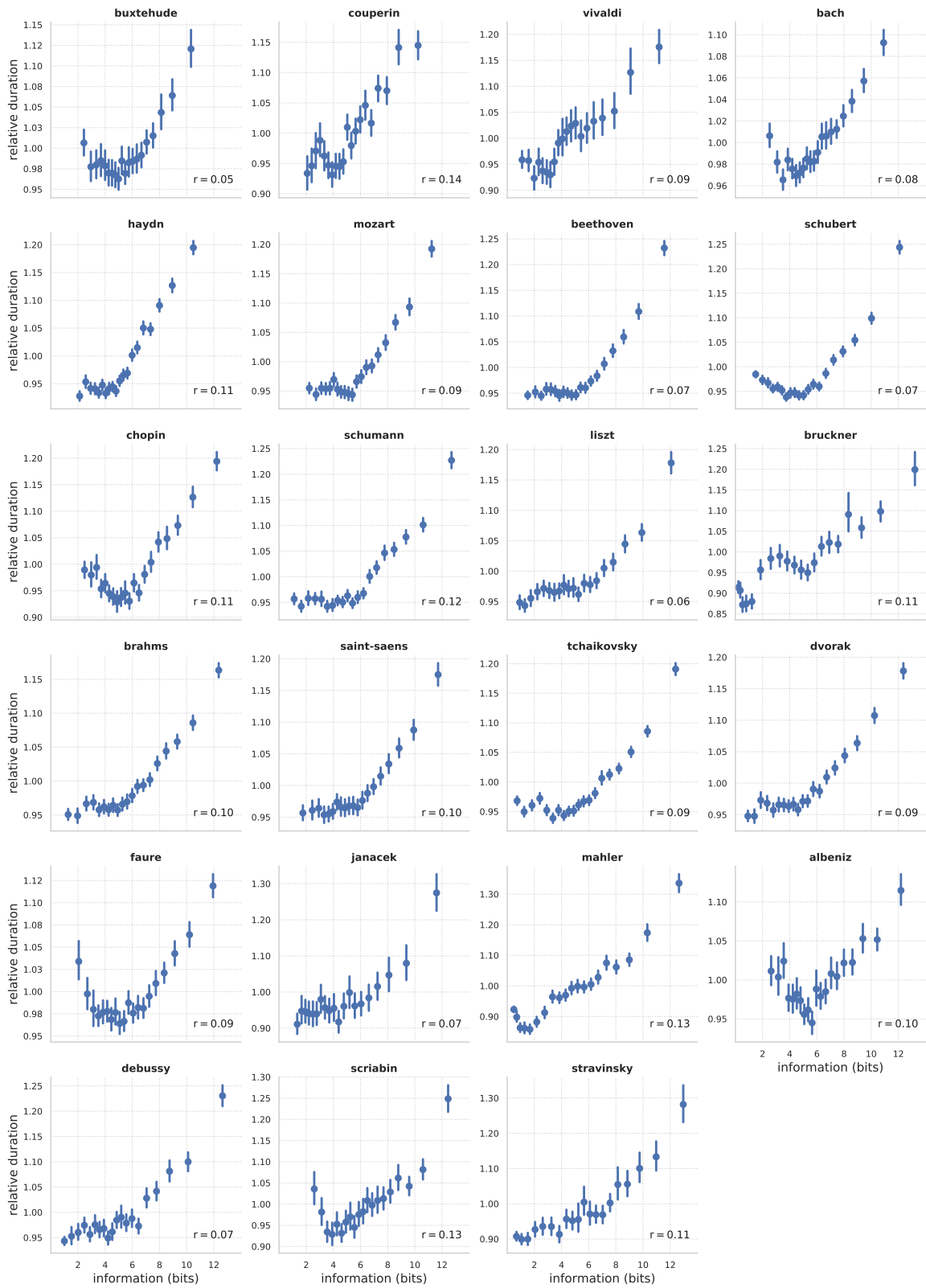


Figure S2: Relative duration as a function of information content, for each composer in the sequenced corpus.

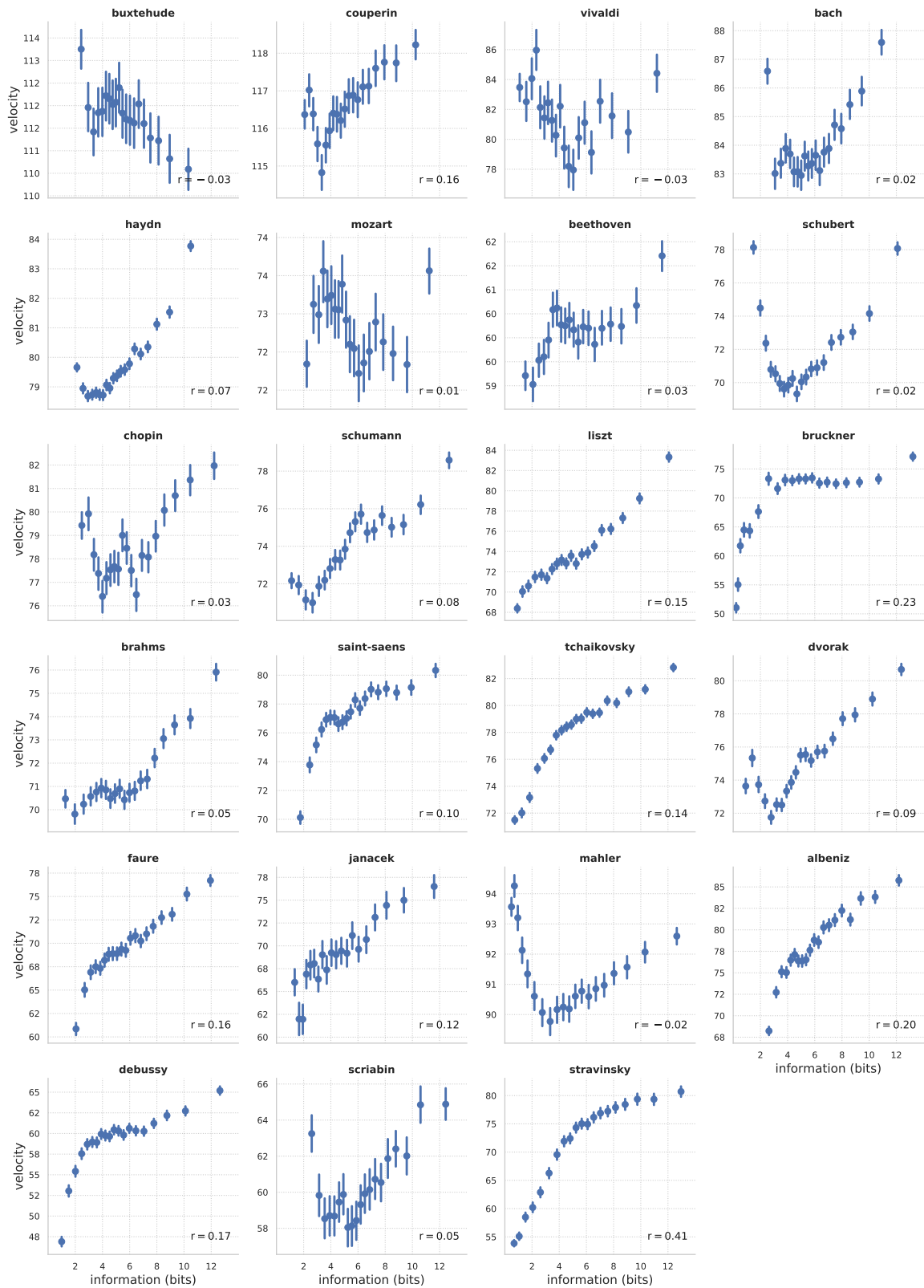


Figure S3: Absolute MIDI velocity, proportional to the loudness, as a function of information content, for each composer in the sequenced corpus.

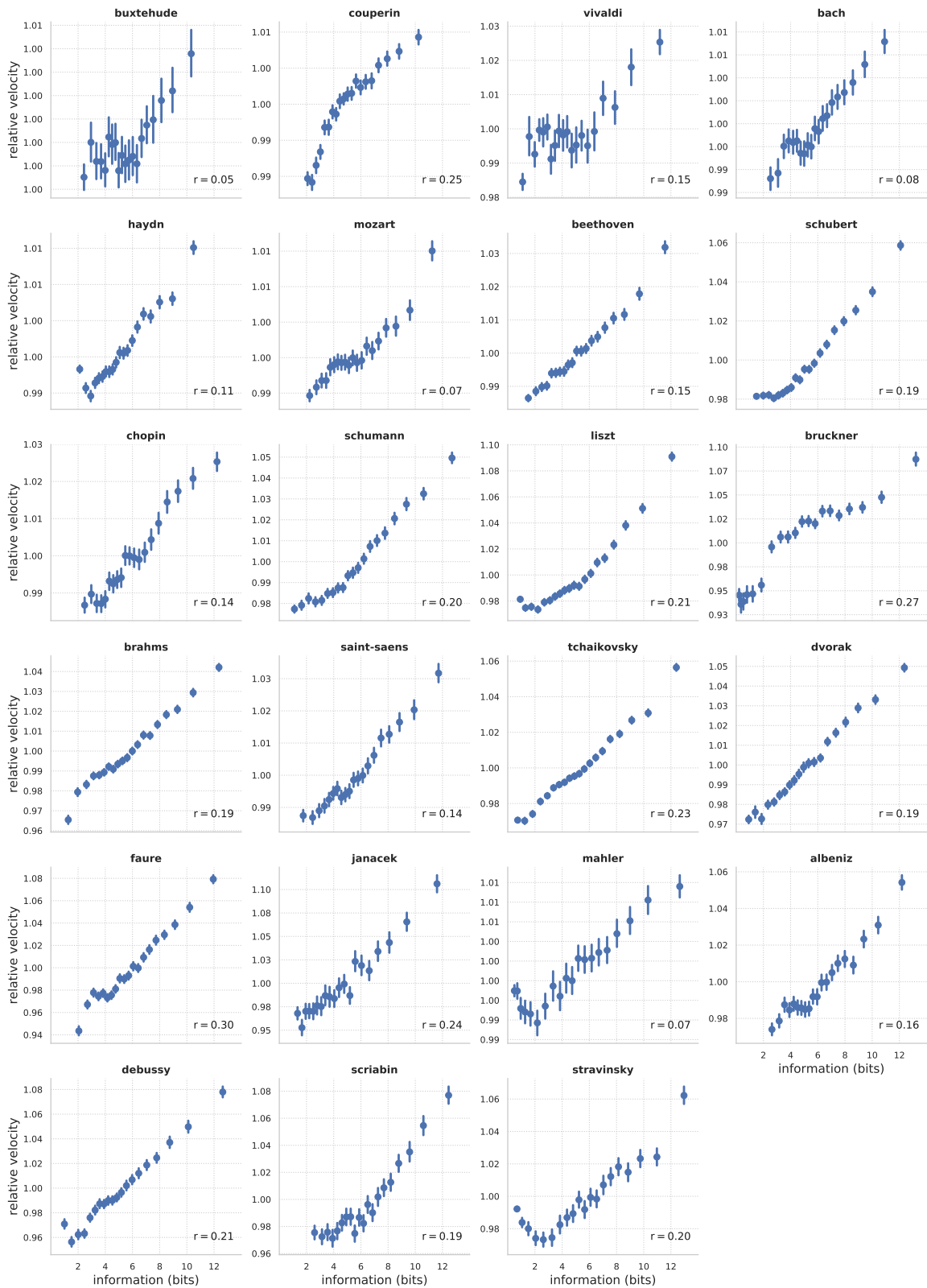


Figure S4: Relative velocity as a function of information content, for each composer in the sequenced corpus.

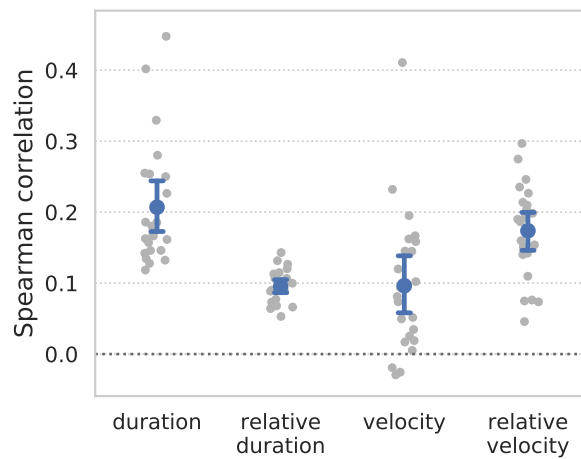


Figure S5: **Distribution of the Spearman correlation coefficients, computed for each composer in the sequenced corpus, between information content and the four variables of interest considered in this study, namely duration, relative duration, velocity and relative velocity.** In each case, the grey dots correspond to individual composers, while the blue dot represents the mean. Error bars indicate the 95% confidence intervals, estimated by bootstrap.

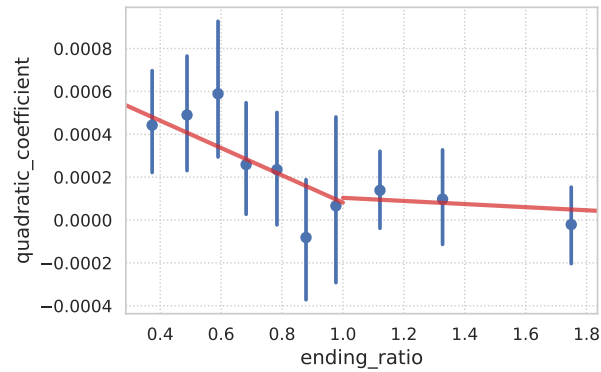


Figure S6: Analysis of the ending of each piece, or why low information content can be associated with greater duration. Looking at the ending (the last 50 segments) of each MIDI file of the sequenced corpus: U-shapeness as a function of the decrease in information. The U-shapeness score is evaluated as the quadratic coefficient estimated on the whole corresponding piece in order to evaluate the U-shapeness of the relationship between information and duration, while the decrease in information toward the end of each piece is evaluated as the ratio between average information of the last two values and the average information of information between positions -50 and -30. Values of ending ratio below 1 indicates a decrease in information, and the lower the value the sharper the decrease. For values of this ending ratio below 1, there is a negative relationship with the U-shapeness score ($r = -0.10$, $p = 2.9e-4$, $z = -3.6$): the stronger the decline in information, the more U-shaped the relation between duration and information is. The red line depicts a piecewise linear fit.

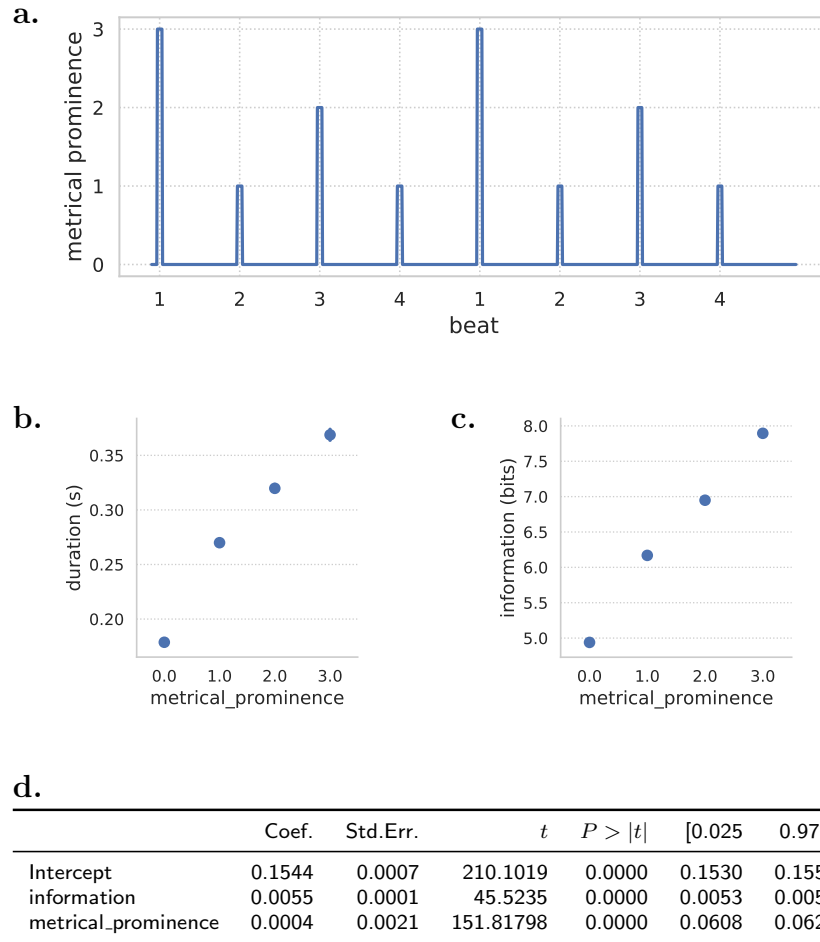


Figure S7: **Analysis of a possible confound: musical meter, considering all pieces with a time signature of 4/4 in the written corpus.** (a) Our model of metrical prominence as a function of the position of the onset of a given segment within a bar. (b) Duration as a function of metrical prominence. (c) Information content as a function of metrical prominence. (d) Results of the linear model ‘duration \sim information + metrical_prominence’ (output from the statsmodels python package).

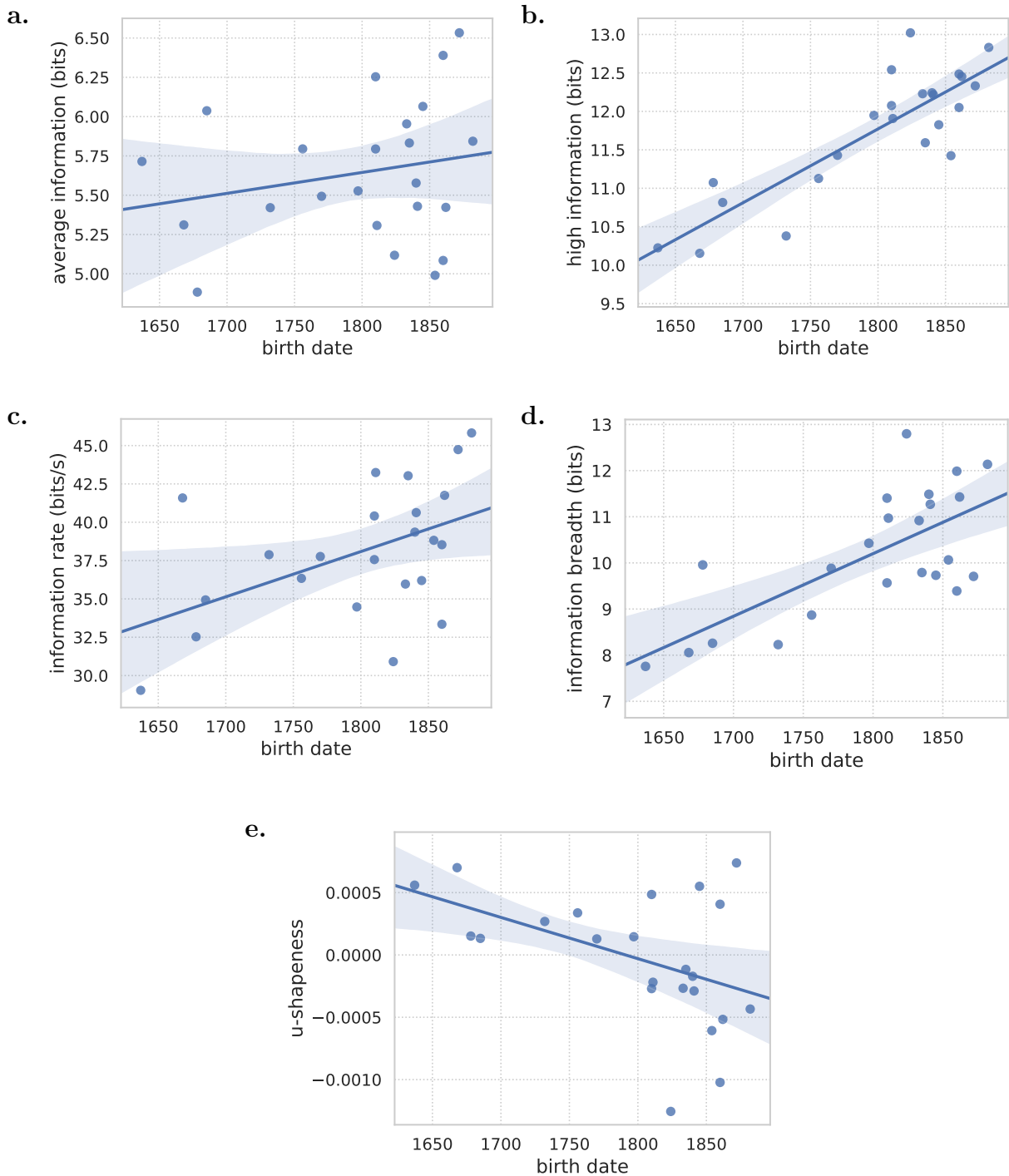


Figure S8: **Diachronic changes of information-related quantities.** As a function of the birth date of the composer, (a) average information ($r = 0.22$, $p = 0.32$, $z = 1.0$), (b) high information, defined as the 95th percentile ($r = 0.72$, $p = 1.2e-4$, $z = 4.7$), (c) information rate, information per unit of time, ($r = 0.48$, $p = 0.019$, $z = 2.5$), (d) information breadth, defined as the difference between the 95th and the 5th percentile ($r = 0.56$, $p = 5.1e-3$, $z = 3.1$), and (e) U-shapeness score, estimated as the quadratic coefficient of a second order polynomial fit ($r = -0.42$, $p = 0.048$, $z = -2.1$). Correlations are Spearman's.