Quantifying Information Structure in Sperm-Whale Codas

Pål Eivind Storvik Olsen

Abstract

Recent work in Nature Communications argues that sperm-whale codas have a combinatorial structure: a core, context-independent form defined by rhythm (the pattern of click timing after normalizing for duration) and tempo (overall duration), and two context-dependent delivery features: rubato (small duration drift) and ornamentation (a terminal extra click). Building on that framework, I analyze 3,673 codas and define the symbol ("unit") as (rhythm type, tempo type) with 18 rhythm templates and 5 tempo bins (18×5 = 90 possible types). Using ornament-free timings for both rhythm and tempo, and defining rubato as the duration change to the immediately previous coda by the same whale with the same unit within an exchange, I estimate marginal entropy, conditional entropies, and the order-1 entropy rate with an exchange-preserving shuffle null. Results show H(Unit) = 3.490 bits, an order-1 rate of 2.079 bits (well below the within-exchange shuffle null; mean 2.549, 95% [2.517, 2.581]), and small but reliable contextual effects: I(Unit; rubato) = 0.073 bits; I(Unit; ornament) = 0.038 bits; joint 0.101 bits. These findings indicate strong short-range sequential structure, modest contextual modulation, and statistical dependence between the contextual and non-contextual coda features.

Acknowledgements

This project builds on Sharma, P., Gero, S., Payne, R., Gruber, D. F., Rus, D., Torralba, A., & Andreas, J. (2024). Contextual and combinatorial structure in sperm whale vocalisations. Nature Communications. The paper introduces the rhythm/tempo (context-independent) vs. rubato/ornamentation (context-dependent) framework on which my unit definition and analyses are based. The annotated coda material originates from the Dominica Sperm Whale Project (DSWP), as described by the authors. (Article is open-access under CC BY 4.0; credit retained.)

Background

The Nature study formalizes coda structure as a combination of rhythm (shape in normalized time) and tempo (overall duration) as context-independent features, with rubato (local, context-driven duration drift between adjacent codas of the same rhythm and tempo by the same whale) and ornamentation (a terminal extra click) as context-dependent features. This motivates separating core form (rhythm×tempo) from delivery (rubato, ornamentation) when quantifying information content and sequence structure.

Data and Unit Definition

– Dataset: 3,673 codas, sorted by exchange (conversation) and timestamp.
– Unit: (rhythm_type, tempo_type) with 18 rhythm templates × 5 tempo bins = 90 possible types.
– Rhythm assignment: template matching against normalized cumulative click-time profiles, computed without terminal ornaments.
– Tempo assignment: nearest tempo center using ornament-free durations.
– Context variables: ornamentation (binary terminal click); rubato (duration change relative to the immediately previous coda by the same whale with the same unit, within the same exchange), binned into {−1 decreasing, 0 stable, +1 increasing}. Codas without a valid predecessor are excluded from rubato-conditioned computations.

Methods

Entropy is "average unpredictability," measured in bits (1 bit = a fair coin flip). H(Unit) tells us how many bits are needed, on average, to encode one coda without using any context. Conditional entropy (e.g., H(Unit|rubato)) asks the same question within each context bucket (here:

decreasing/stable/increasing). The order-1 entropy rate, $H(U_t \mid U_{t-1})$, tells us how much uncertainty remains about the next coda once we know the previous one, a direct measure of short-range structure.

I estimate H(Unit) with a standard plug-in estimator, and compute H(Unit|rubato), H(Unit|ornament), and H(Unit|rubato,ornament) on the appropriate subsets (rubato analyses restricted to codas with valid rubato). Mutual information is I(Unit; X) = H(Unit) − H(Unit|X) on matching samples. For sequence structure, I compute $H(U_t|U_{t-1})$ within exchanges only (no cross-exchange transitions) and compare it to a null that shuffles codas within each exchange (preserving exchange length and per-unit counts) before recomputing the rate; I report the null mean and its 95% interval.

Results

Marginal and conditional entropies

- H(Unit) = 3.490 bits (full sample).
- On the full sample: H(Unit|ornamentation) = 3.452 bits ⇒ I(Unit; ornamentation) = 0.038 bits (~1.1% of the full baseline).
- On the rubato-valid subset: H(Unit) = 2.962 bits; H(Unit|rubato) = 2.889 bits ⇒ I(Unit; rubato) = 0.073 bits (~2.5% of the subset baseline).
- Jointly, H(Unit|rubato, ornamentation) = 2.861 bits ⇒ I(Unit; rubato, ornamentation) = 0.101 bits (~3.4% of the subset baseline).

Sequential structure

The observed order-1 entropy rate is 2.079 bits. Under the within-exchange shuffle null, the rate is 2.549 on average, with a 95% interval [2.517, 2.581]. The observed rate is ~0.47 bits below the null band, indicating robust short-range sequential structure: knowing the previous unit reduces uncertainty about the next by about 1.41 bits relative to the marginal (3.490 − 2.079), a ~40% reduction.

Throughput

Mean ornament-free coda duration is 0.91 s. Average information per second (marginal) is 3.82 bits/s (3.49 / 0.91). For context, the predictive throughput implied by the rate is ~2.28 bits/s (2.079 / 0.91).

Discussion

This analysis supports two complementary conclusions about sperm-whale codas. First, there is clear short-range sequential structure. The order-1 entropy rate $H(U_t|U_{t-1})$is 2.079 bits, well below both the marginal entropy H(Unit) = 3.490 bits and a within-exchange shuffle null (mean 2.549; 95% [2.517, 2.581]). Practically, knowing the previous coda reduces uncertainty about the next by about 1.41 bits (≈40%). Because the null keeps each exchange's length and unit counts the same while randomizing their order, the gap to the null isolates true ordering rules rather than differences in how often each unit appears.

Second, rubato and ornamentation are not statistically independent of the rhythm×tempo unit. On the rubato-valid subset, H(Unit) = 2.962 bits and H(Unit|rubato) = 2.889 bits, giving I(Unit; rubato) = 0.073 bits; on the full sample, H(Unit|ornamentation) = 3.452 bits, giving I(Unit; ornamentation) = 0.038 bits. These values are small in absolute terms, roughly 1–3% of their respective baselines, but positive. Intuitively, delivery features (slight local timing drift and the presence/absence of a terminal click) carry weak but real cues about which rhythm×tempo unit is used. The joint value I(Unit; rubato, ornamentation) = 0.101 bits is close to the sum of the two individual effects, suggesting near-additivity with little redundancy or synergy. Mutual information quantifies association, not causation, so the takeaway is a modest statistical coupling between form (rhythm×tempo) and delivery (rubato/ornament).

These results fit the combinatorial picture advanced in recent work: whales appear to draw from a core inventory of forms (rhythm×tempo), add context dependent features, and arrange codas in locally structured sequences. The contrast between capacity (e.g., $\log_2 90 = 6.49$ bits/coda for rhythm×tempo under a uniform, independent model) and realized information (3.49 bits marginally; 2.08 bits at order-1) shows that while the coding space is large, actual usage is non-uniform and dependent, not uniform and memoryless. That pattern is exactly what we would expect in a communicative system shaped by social interaction and behavioral context.

Limitations

Rubato labels exist only where a same-whale, same-unit predecessor is available; rubato-conditioned metrics are therefore computed on a subset. I report point estimates and a dependence-respecting null for the entropy rate (rather than confidence intervals) because exchange-level dependence and variable exchange sizes would require more elaborate resampling. The threshold used to bin rubato (decreasing / stable / increasing) was chosen manually (guided by the Supplementary figures); different thresholds would shift the exact numbers, though the overall pattern is stable in small sensitivity checks.

Reproducibility

All analyses use fixed random seeds; shuffles occur within exchanges; environment versions are pinned; figures and a run-metadata file (seed, package versions, counts, headline numbers) are written automatically by a single entry point. Average information per second uses ornament-free durations to match the unit definition.

Conclusion

Under a rhythm×tempo unitization inspired by the Nature framework, sperm-whale codas carry ~3.49 bits per coda marginally and ~2.08 bits per coda at order-1, with the observed rate far below a within-exchange shuffle null, evidence of strong short-range sequential structure. Rubato and ornamentation add small but consistent contextual information, supporting a combinatorial system in which delivery features modulate a predictable core form. The contrast between capacity and realized information highlights that while the coding space is large, natural usage is structured and non-uniform.

References

Sharma, P. et al. (2024). Contextual and combinatorial structure in sperm whale vocalisations. Nature Communications. DOI: 10.1038/s41467-024-47221-8.
Sharma, P. sw-combinatoriality (code/data repository). GitHub