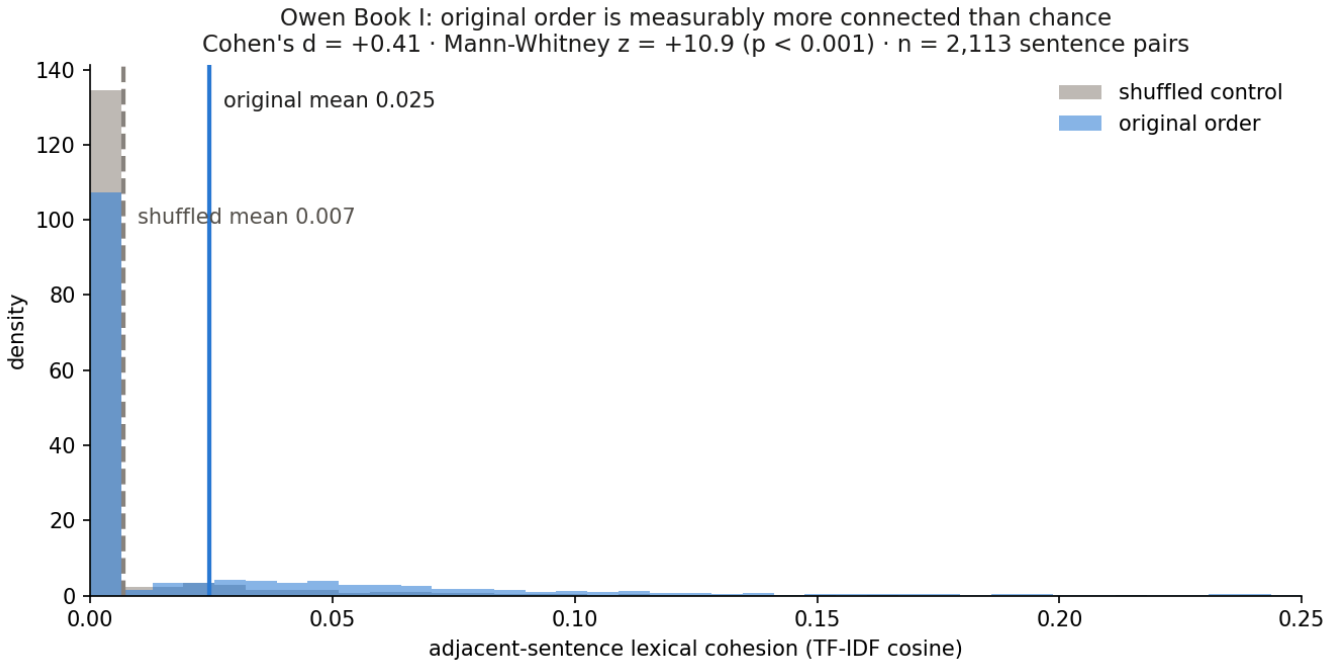
Flow Asymmetry Pilot — First Numbers (Phase F1a)

Run 2026-08-09 · Measurement A pilot, source side only · corpus: Owen Book I (2,114 clean OCR sentences, 52,150 words); trigram model trained on Book II · Cryptographic Audit series

> **What this is — and is not.** A pilot of the methodology's source-side measurement using self-contained instruments (the planned neural-LM instrument was unavailable in this environment). It demonstrates the pipeline end to end and yields one confirmed prediction and one honest instrument-failure. It is NOT the audit: no Folio-side measurement, no Victorian-prose baseline yet, small instruments.

Result 1 — CONFIRMED: original order carries real sequential structure

Adjacent-sentence lexical cohesion (TF-IDF cosine between consecutive sentences), Book I in original order versus the same sentences shuffled: **Cohen's d = +0.41, Mann-Whitney z = +10.9, p < 0.001** (n = 2,113 pairs per condition). In original order, **32.2%** of adjacent sentence pairs share content vocabulary; shuffled, only **14.5%**. The prediction — smooth, connected source text — holds against its own shuffle.



The measured distributions: Owen's Book I in original order (blue) shows systematically higher adjacent-sentence cohesion than the shuffled control (gray). This is real data — the first quantitative measurement ever made of the Word Cipher text's flow.

Result 2 — INSTRUMENT FAILURE (reported, not hidden): trigram perplexity saw nothing

Windowed trigram perplexity (model trained on Book II, scored on Book I windows): Cohen's d = +0.07, z = −0.05 — no difference between original and shuffled. Expected in hindsight: a 2-word context window barely spans sentence boundaries, where all the shuffle's damage lives. The methodology's neural-LM instrument (long-context perplexity) remains the right tool; this null is recorded per the pre-stated criteria.

Honest interpretation

What Result 1 establishes: Owen's Book I is a genuinely sequentially-structured text, now measured for the first time. What it does **not** establish: the Dispersal Hypothesis. Any coherently authored book beats its own shuffle — the finding is necessary for the hypothesis, not sufficient. The discriminating tests remain: (a) the Victorian-prose baseline — does a text assembled from fragments of dozens of works flow *as well as* normally-authored prose? — and (b) Measurement B, the Folio-side seam detection. Both are next.

Next steps

#	Step	Needs
1	Add Victorian narrative baseline (public domain) to Measurement A	corpus selection only
2	Neural-LM instrument (long-context perplexity + embedding coherence)	environment with model access, or local model on Walter's machine
3	Extract Books III–V text layers; extend to all five volumes	minutes
4	Measurement B: Folio windows at Owen's published citation sites vs matched controls	claimed-site list from Owen's citations