FILED FRONTMATTER

```
title: "The Leverage Inversion: Dating the Transition from Consuming AI
Answers to Directing an AI Workforce in a 3.5-Year Single-Subject
Conversational Corpus"
author: "Paul Roebuck"
project: "Corpus_Expeditions_2026-07-10"
created: 2026-07-16
type: preprint
doc: "05-e"
version: "v0.1"
tags: [shads, corpus-expeditions, leverage-inversion, preprint]
aliases:
  - "Leverage Inversion preprint"
```

The Leverage Inversion: Dating the Transition from Consuming AI Answers to Directing an AI Workforce in a 3.5-Year Single-Subject Conversational Corpus

Paul Roebuck — independent practitioner-researcher, UK · paulsroebuck@gmail.com

Preprint v0.1 · 16 July 2026 · Not peer reviewed. Posted at paulroebuck.co.uk/blogai and minditapparatus.netlify.app/preprint. Comments welcome.

Abstract

This is an autoethnographic single-case study: the author is the subject, and the data are the author's complete human–AI interaction corpus — 2,286 threads, 47,442 turns, 986,823 subject-typed words and 6,865,588 AI prose words across ChatGPT, claude.ai and Claude Code, spanning 31 December 2022 to 10 July 2026. From it we date and characterise a regime change we call the *leverage inversion*: the transition from using AI primarily as an answer engine to directing it as a collaborative workforce. The inversion is not, as naively expected, the AI-to-human word ratio falling because the human asks for less; the ratio falls (19.2:1 at the December 2025 peak to 3.75:1 integrated by July 2026) because the human's own output explodes, from 4–21k words/month through H2 2025 to 105–163k words/month from April 2026. The naive question-to-directive grammar shift

is falsified in this corpus: question share is trendless (16–29% throughout), and directive-by-verb share actually halves into a 2025 trough. The true leading indicator is the rise of *declarative steering* — messages that supply context, judge output, or correct course — from 44.5% of the subject's messages in 2023-Q4 to 63.6% in 2025-Q4, before any volume change. The inversion proper is a ~6-week transition (1 April – 19 May 2026), with onset in an explicit workforce-design week (13–20 April 2026) and consummation on 10 May 2026, when the first AI thread was treated as a named employee with a retirement. A full dress-rehearsal 11 months earlier (May–June 2025) showed the complete production signature and then aborted, indicating that capability access alone is insufficient: a forcing project is required. We state the pattern as a testable three-stage model with a metrics kit runnable on any user's chat export.

1. Introduction

What is known about how people use conversational AI comes overwhelmingly from population-scale cross-sections. OpenAI's analysis of 1.1 million ChatGPT conversations maps what messages ask for at a point in time [8]; Anthropic's Economic Index maps millions of Claude conversations onto occupational task categories [7]. These studies answer *what the population does*. They cannot, by design, answer a different question: *how does one sustained user's relationship with these systems reorganise over years?* Longitudinal, within-person evidence is nearly absent — few users keep their complete record, and platform exports were not designed for research.

This paper offers one such record. The author has used conversational AI continuously since 31 December 2022 and holds complete exports of all three platform legs. The corpus captures, in one person, the period in which large language models went from novelty to infrastructure — and it captures a specific, datable behavioural regime change that we argue is the interesting unit of analysis: the moment the human stopped buying answers and started directing a workforce. Licklider's founding vision of man–computer symbiosis [1] imagined the human setting goals and the machine doing the routinised work; this corpus records, at the scale of one working life, the months in which that division of labour actually arrived — and shows that what changed first was not the machine but the human's grammar.

Methodologically this is an autoethnographic single-case study [3, 4], a genre with precedent in human–computer interaction [5]: the author is simultaneously investigator and subject. We state this openly and design around it (Section 7). The approach trades generality for a kind of evidence no other design can produce: complete, timestamped, behavioural (not self-reported) coverage of a single human–AI relationship over 3.5 years. That corpus-based signals can reliably detect AI's influence on human behaviour is

established at population scale [6]; we apply the same logic within-person. The contributions are: (i) a dated, quantified account of the leverage inversion; (ii) the falsification of an intuitive indicator and its replacement with a better one; (iii) a three-stage model stated with falsifiable orderings; and (iv) a portable metrics kit computable from any user's own export.

2. Data and ethics

Three export legs, parsed to a common per-conversation/per-session schema:

Leg	Threads	Span	Subject-typed words	AI prose words
ChatGPT	1,928	31 Dec 2022 – 9 Jul 2026	586,208	4,097,618
claude.ai	279	3 Oct 2023 – 9 Jul 2026	219,849	1,754,361
Claude Code	79 sessions (+239 subagent transcripts)	19 May 2026 – 10 Jul 2026	180,766	668,368 (+345,241 subagent)

Plus 689,643 words of subject-supplied attachments on the claude.ai leg. Claude Code tool traffic (5.87M words of tool inputs/results) is excluded from all ratios; "AI words" means prose addressed to the subject. All subject-typed messages (N = 20,060) were classified by speech-act class in a fresh scan; scripts and intermediate tables are retained in the project archive (see Data availability).

Ethics. The author is the sole human subject and consents by construction. The underlying corpus includes professionally sensitive material (the author practised as a psychotherapist during the early corpus years); that material is governed by a separate data-governance charter, contributes only to aggregate word and message counts here, and is not quoted, described, or individually identifiable in this paper or its derived tables. No third party's conversational content appears in any output. The raw corpus is not shareable; derived aggregates are (Data availability).

3. Metrics

Six measured series, all monthly unless stated:

- **Leverage ratio R(t)** — AI prose words / subject-typed words, integrated across legs.
- **Subject output U(t)** — absolute subject-typed words, all legs.
- **Message length L(t)** — subject words per subject message, by leg.

- **Speech-act mix** — each subject message classified, in the spirit of speech-act theory [2], by a lexical classifier as *question* (leading interrogative token or terminal "?"), *directive* (leading imperative verb, ~250-verb list), *assent* (leading yes/ok/thanks-class token), or *other* = **declarative steering** (context supply, judgement, correction, continuation cues). Reported quarterly; N = 20,060.
- **Thread depth D(t)** — turns per conversation (mean/median/max) by leg.
- **Production-structure markers** — attachment words per month (artefact supply); named-seat births per month (custom-titled working instances, including a "— Retired" naming convention for completed ones).

4. The inversion, dated

4.1 The phase sequence

The corpus divides cleanly into five phases:

Phase	Period	Signature
1. Consumption	Dec 2022 – Apr 2025	Median depth 4–8 turns; directive-by-verb at lifetime high (content-generation commands); R mostly 2–11
2a. Failed first inversion	May – Jun 2025	An app-build project: 49 threads, 330,091 attachment words, U spikes 4× — then aborts; the claude.ai leg goes silent for 9 months
2b. Peak consumption	Jul – Dec 2025	R climbs 14.9 to **19.2 (Dec 2025, all-time peak)**; U flat at 4–15k; L flat at 20–25 words
3. Onset	Jan – Apr 2026	January ratio half-step (19.2 to 7.0, topics still consumer); 13–20 Apr: a workforce is explicitly designed; U hits 105,662 (7.3× Dec)
4. Consummation	10 – 19 May 2026	First named-employee thread born 10 May (with a hire date and, eventually, a retirement); a five-instance succession follows 24–29 May; first Claude Code session 19 May
5. Operation	Jun – Jul 2026	29 named seats born in June; U (Claude Code alone) = 100,706 in June; integrated R = 3.75 by Jul; ChatGPT reverts to a 4–8-turn errand desk

4.2 What actually inverted

The inversion is in the **denominator**. AI output kept growing through the transition (1.31M words in May 2026, the largest month ever); the ratio fell because subject output grew faster. Three co-movements confirm a regime change rather than a mere intensity shift:

- **L(t)**: 20.5–25.0 words/message across H2 2025 (ChatGPT) vs 106.9 (claude.ai, Apr 2026) and 109.4 (Claude Code, Jul 2026) — query length became briefing length.
- **D(t)**: ChatGPT median depth is 4–8 turns in essentially every month 2023–2026; Claude Code sessions run mean 84–125 turns (max 1,161). Depth exploded only on the production estate.
- **Structure**: zero named seats anywhere in the corpus before 10 May 2026; 44 in the 62 days after. Attachment supply, absent for 8 straight months, restarts March 2026 and runs 34–147k words/month thereafter.

4.3 Leading vs lagging indicators

- **Leading (moves through 2025, before any volume change):** declarative-steering share rises every quarter of 2025 — 50.6%, 52.2%, 58.9%, 63.6% — the subject progressively tells, shows and judges rather than asks. A critical vocabulary about AI failure modes is also built inside peak consumption (first "hallucinating" 16 Apr 2025; a contested sycophancy episode 8 Sep 2025).
- **Falsified as an indicator:** the question-to-directive grammar shift. Question share is trendless (0.16–0.29 band across 3.5 years); directive-by-verb share FALLS from a 2023-Q3 peak (69% of classified question+directive mass) to a 2025-Q3/Q4 trough (29%), recovering only partially in 2026. Judgement: 2023–24 directives were vending-machine content commands ("write/draw/summarise X"); the grammar class recurs in 2026 as work-orders to named workers — same syntax, different social relation, which is why grammar alone cannot date the inversion.
- **Coincident:** subject output U(t) (7.3× step in April 2026) and message length L(t).
- **Lagging:** the ratio fall itself, and infrastructure adoption (the agentic-tooling leg begins 19 May, nine days AFTER the consummation date).

4.4 The boundary, ruled

Onset: 13–20 April 2026 — a workforce-design week (three consecutive threads explicitly designing AI co-worker roles and preferences, and the coining of a working term for the arrangement) inside the first exploded-output month. **Consummation: 10 May 2026** — the first AI thread treated as a named employee with a hire date and a retirement; the single date if one is forced. The whole transition spans ~6 weeks (1 April – 19 May 2026).

5. Mechanism

Three candidate causes, tested against ordering in the data (interpretive readings labelled *judgement*):

- **A forcing project — supported, as consolidator.** The May–June 2025 episode proves capability and competence were present 11 months early: full production signature (module decomposition, 330k attachment words, versioned builds), then abort and reversion to the steepest consumption climb in the corpus. What differed in April–May 2026 was a project with a deadline-shaped identity payoff (a book), which arrived weeks after onset and locked the behaviour in. *Judgement:* a forcing project is necessary to make the inversion stick; it is not what starts the drift.
- **Supply side (better long-output models from late 2025) — explains the peak, not the inversion.** Long-answer models stretch the numerator, producing the 19.2:1 December 2025 peak. But the inversion generalised across platforms: the ChatGPT-leg ratio fell to 4.1–9.1 through 2026 even as ChatGPT thread count tripled, with depth unchanged — the old platform was re-purposed, not replaced by a better one. A pure supply-side story predicts the opposite (ratio keeps climbing wherever output is cheapest).
- **Demand side (the practitioner's stance) — supported as the slow variable.** The declarative-steering rise through 2025 is a 12-month behavioural drift that precedes tools, book and vocabulary. *Judgement:* this is a therapist-coach instinct — direct, judge, supervise — progressively applied to the machine; day zero itself (31 December 2022, the subject feeding his own article in for critique) was a supply act, and a 2025 photography-critique project rehearsed the feed-material-judge-output loop at scale.

Synthesis (judgement): a slow demand-side drift (steering share rising) + a latent capability proven in a failed trial + a forcing project = a fast (~6-week) phase change. The order matters: stance moved first, volume second, infrastructure third, ratio last.

6. Generalisability

6.1 The three-stage claim (testable)

For a sustained individual user of conversational AI:

- **Stage 1 — Consumption.** R(t) trends upward; U(t) flat; L(t) low and flat; shallow threads; no artefact supply. The user is buying answers.
- **Stage 2 — Saturation.** R(t) reaches its lifetime maximum while U(t) stays flat; the leading indicator is composition, not volume — declarative-steering share of user messages rises monotonically; critical vocabulary about AI failure modes appears; zero or more *failed trial inversions* (production-signature bursts that abort) may occur.
- **Stage 3 — Production inversion.** U(t) steps up by >3× trailing median and holds; L(t) at least doubles versus the Stage-2 floor; R(t) falls substantially from peak *while AI output does not fall proportionally*; thread depth bifurcates (a deep production estate

plus a shallow errand desk); delegation structure appears (named/roled threads, artefact supply, succession conventions).

Falsifiable orderings: steering share rises BEFORE the U(t) step; the R(t) peak precedes the inversion; infrastructure adoption follows rather than leads the behavioural change. Any subject showing the U(t) step without the prior steering rise, or an R(t) fall driven by AI output collapse rather than user output growth, falsifies the model for that case.

6.2 Metrics kit

Computable from any ChatGPT/Claude export pair: monthly R(t), U(t), L(t) per leg; speech-act mix by the lexical classifier (question / directive / assent / declarative-steering); thread-depth distribution per leg; attachment words; naming-convention scan over titles. Declare Stage 3 when, for 2+ consecutive months: U(t) > 3× trailing-12-month median AND R(t) < 0.6× lifetime peak AND AI words ≥ 0.5× their own trailing median. Date onset at the first month of the U(t) step; date consummation at the first delegation-structure marker.

6.3 Bridge note

The consumption-to-production curve is a candidate behavioural correlate for human-side disposition instruments the author is developing separately: Stage 1 consumes what the field offers, Stage 3 directs it. We note the bridge and deliberately do not build it here; testing the mapping would require instrument scores alongside export metrics for multiple subjects.

7. Limitations

- **Single subject, and the author.** One person's corpus, and an unusual one (therapist-coach, later building an AI-behaviour framework and a book on exactly this material). The three-stage model is generated from this case, not yet tested on any other. The author's dual role as investigator and subject is declared, standard for the autoethnographic genre [3, 5], and partially mitigated by the design: every claim rests on timestamped behavioural data computed by scripted scans, not on recollection or self-report.
- **Reactivity (observer effect).** The subject's deliberate self-study of this corpus began in mid-June 2026. The inversion window (April–May 2026) predates it, so the central dating is not an artefact of self-observation; Phase-5 (June–July 2026) measurements, however, describe a subject who knew he was measuring himself.
- **Export scopes differ by leg.** ChatGPT has no attachment field (pasted material counts as typed words, inflating U and L on that leg); Claude Code has no cloud export and its

counts exclude tool traffic by construction; one agentic product leg used in the period is absent from all exports, so named-seat counts and 2026 volumes are floors.

- **Survivorship.** Deleted conversations are invisible; the corpus is what survived to export day (10 July 2026). July 2026 covers 10 days.
- **Classifier is lexical.** Leading-token plus terminal-"?" rules; the declarative-steering class (45–65% of messages) is heterogeneous and was sub-typed only by qualitative sampling. Hand-coded calibration of the classifier is planned and will be reported in a revision. Word counts are whitespace-based; timestamps UTC.
- **Reproducibility note.** The first-generation scan script was accidentally overwritten during the completion pass behind this analysis; the script now archived is a documented reconstruction verified to reproduce the published quarterly shares within ~1 percentage point (population delta 2.4%). The first-generation intermediate tables are intact and remain authoritative.
- **Ratio definition.** R(t) uses AI prose only; including reasoning/tool traffic would raise 2026 numerators substantially and is a different (machine-effort, not communication) measure.

Data and code availability

The raw conversational corpus cannot be shared (it contains professionally sensitive and third-party material; see Ethics). The derived monthly and quarterly aggregate tables and the portable scan scripts behind every figure in this paper are retained in a versioned project archive and are available from the author on reasonable request; deposit in a public repository (with a DOI) is planned and this preprint will be updated with the link.

Competing interests

The author holds a pending UK trade mark application for an AI-behaviour framework (SHaDS™) developed from this corpus, is the author of a book drawing on the same material, and has commercial interests in related assessment instruments. This paper reports behavioural measurements only and does not describe or depend on any proprietary instrument.

AI-assistance statement

The corpus scans, metric computations and a first analytical draft were produced by Claude-based research agents (Anthropic) working under the author's direction and brief on 10 July 2026, with the working attribution "The Excavator"; this manuscript was revised for public issue with AI assistance under the author's editorial control. The author reviewed the analyses, rules on all interpretive judgements, and takes sole responsibility

for the content. The reflexive wrinkle is acknowledged: the instruments used to study the human–AI relationship are of the same kind as the systems being studied.

References

- [1] Licklider, J. C. R. (1960). Man-Computer Symbiosis. *IRE Transactions on Human Factors in Electronics*, HFE-1, 4–11.
- [2] Searle, J. R. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press.
- [3] Ellis, C., Adams, T. E., & Bochner, A. P. (2011). Autoethnography: An Overview. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 12(1), Art. 10.
- [4] Yin, R. K. (2018). *Case Study Research and Applications: Design and Methods* (6th ed.). Sage.
- [5] Lucero, A. (2018). Living Without a Mobile Phone: An Autoethnography. *Proceedings of the 2018 ACM Designing Interactive Systems Conference (DIS '18)*. doi:10.1145/3196709.3196731
- [6] Kobak, D., González-Márquez, R., Horvát, E.-Á., & Lause, J. (2025). Delving into LLM-assisted writing in biomedical publications through excess vocabulary. *Science Advances*, 11, eadt3813. doi:10.1126/sciadv.adt3813
- [7] Handa, K., Tamkin, A., et al. (2025). Which Economic Tasks are Performed with AI? Evidence from Millions of Claude Conversations. arXiv:2503.04761.
- [8] Chatterji, A., Cunningham, T., Deming, D. J., Hitzig, Z., Ong, C., Shan, C. Y., & Wadman, K. (2025). How People Use ChatGPT. NBER Working Paper 34255.

Preprint v0.1 · Paul Roebuck · 16 July 2026 · derived from Expedition 05, Corpus_Expeditions_2026-07-10 (internal archive). Reference verification: items [5]–[8] checked against their public records on 16 July 2026.

The Tomb Map →