

Swift-Footed Autocomplete: The Homeric Noun–Epithet System as a Zero-Click Predictive-Text Engine, with a Defect Audit of Its Longest-Running Production Incident

Philomena Q. Hexworth-Adeyinka¹ Telemachus J. Braithwaite-Osei² Dagmar V. Papastavrou-Lindqvist³

¹Chair of Computational Philology, Marlowe Institute

²Laboratory for Oral Tradition Engineering, Kelvin College

³Centre for Bronze Age Human–Computer Interaction, Ossory Institute

Abstract. Predictive text—the phone’s habit of guessing your next word—is widely regarded as a modern convenience that fails often enough to be a genre of joke. We show that it is instead an ancient technology that once worked flawlessly, and that its finest deployment has been in continuous production for roughly 2,750 years. Building on Milman Parry’s observation that the Homeric epithet is selected by the space remaining in the verse rather than by the situation—hence “blameless Aegisthus,” an epithet bestowed on the epic’s principal murderer—we demonstrate that the noun–epithet system of the *Iliad* and *Odyssey* constitutes a *zero-click* completion engine. Conditioned on name, grammatical case, and metrical slot, epithet choice carries 0.032 bits of information: the bard, on average, chooses nothing. We audit the system’s fabled economy as a lookup table and find a 0.9% collision rate; we show in a randomized trial ($N = 48$) that the engine reduces compositional latency 8.7-fold, to near performance speed; and we measure a semantic defect rate of 0.31%, roughly thirty times more reliable than the autocorrect in the reader’s pocket. Modern predictive text fails, we argue, because it guesses meaning inside an unconstrained interface, whereas the hexameter constrained the interface until there was nothing left to guess. Our recommendations include a meter.

Keywords: oral-formulaic theory; predictive text; dactylic hexameter; epithets; human–computer interaction; cache design.

1 Introduction

Every reader of this journal has watched a telephone complete the word “ducking.” The failure is so familiar that it has produced its own folk literature, and the research response has been consistent for two decades: predict harder. Larger models, longer context, more of the user’s past correspondence ingested in the hope of guessing the next word [5, 13]. Acceptance rates for next-word suggestions nevertheless remain near one in five [13], and the direction of travel is understood to be asymptotic.

This paper reports that the problem was solved before the alphabet reached Greece, that the solution ran for roughly a millennium at a defect rate modern vendors would describe as fraudulent, and that we have possessed the complete engineering documentation—in hexameters—since antiquity. The documentation is the *Iliad* and the *Odyssey*.

Our starting point is the oldest result in modern Homeric scholarship. Milman Parry demonstrated in 1928 that the ornamental epithets of Homeric verse—*swift-footed* Achilles, *much-enduring* Odysseus, *rosy-fingered* Dawn—are not selected for meaning. They are selected by the shape of the space remaining in the line [9, 10]. Achilles is swift-footed while seated. The ships are swift while beached, for ten years, growing barnacles. Dawn’s fingers are rosy in all weathers. And in the most discussed line of the *Odyssey*’s first book, Aegisthus—seducer of Clytemnestra, murderer of Agamemnon, the moral nega-

tive example for the entire poem—is introduced as *blameless* Aegisthus, because the adjective *amymōn* is what the meter dispenses for a name of that shape in that position (*Odyssey* 1.29) [10].

Philology has treated this as a fact about poetry. We propose to treat it as a fact about interface design. A system in which the correct completion of an utterance is determined by formal context, requiring no deliberation from the operator, has a modern name: it is the ideal form of predictive text, the state the industry calls *zero-click* interaction, in which the most frequent action happens automatically and the user’s only remaining job is to not object [11]. The industry regards this state as aspirational. We show it shipped in the eighth century BC.

Concretely, this paper makes four contributions:

1. **A measurement of choice.** Across all 27,803 lines of the *Iliad* and *Odyssey*, the information content of epithet selection, conditioned on the metrical context, is 0.032 bits—roughly one coin flip per thirty epithets. Composition is zero-click (§5).
2. **An audit of the cache.** Parry’s “economy”—one epithet per hero per metrical shape—is shown to be a collision-free lookup table to within 0.9%, against a natural-language baseline of 8.4 interchangeable synonyms per key (§6).
3. **A latency trial.** With the formula stock available,

traine rhapsodes compose metrically valid verse 8.7× faster, approaching the real-time benchmark of the South-Slavic singers (§7).

4. **A defect report.** We audit every semantically discordant epithet in both epics (defect rate 0.31%) and formally file the oldest of them, *blameless Aegisthus*, as incident HOM-0001, status WONTFIX (§8).

We close with the obvious implication for the modern industry, which will not enjoy it: the missing component of contemporary predictive text is not a better predictor. It is a meter.

2 Background for the General Reader

We briefly define our terms, as a courtesy both to computer scientists who have not read Homer and to Homerists who have not read their phones. Experts in either field may skip the corresponding paragraph, and are asked to forgive what the other paragraph says about them.

The hexameter. Every line of Homeric epic is a dactylic hexameter: six metrical “feet,” each a dactyl (one long syllable, two shorts: *tum-ti-ti*) or a spondee (two longs: *tum-tum*), with the final foot always two syllables. A line is thus not an open field of language but a small grid of admissible shapes—in practice a handful of standard templates, with a customary breathing-point (the caesura) partitioning the line into reusable halves [14]. The reader should picture less a blank page than a form with six boxes.

The formula. Oral epic was composed *in performance*, without writing, at speaking speed. This is possible because the tradition supplies prefabricated phrases—formulae—each machined to fill one of the grid’s standard shapes exactly [9, 6]. The singer does not recall a text; he assembles the line from stock, live.

The epithet. The most famous formula class pairs a name with a fixed ornament: *swift-footed Achilles*, *ox-eyed queenly Hera*. The ornament’s job is not description but *fit*: each such phrase is a part with a known shape, kept in stock because lines routinely have a hole of that shape [10, 4].

Autocomplete. A modern text interface watches what you type and proposes the next unit, which you accept by tapping or, in the ideal case, by doing nothing. A *cache*, for our purposes, is simply a lookup table: given a key (here: a name, its grammatical case, and the hole to be filled), return the stored value (the phrase). *Latency* is waiting.

The reader now holds the entire technical apparatus of this paper. Our claim, in these terms: the epic tradition built a completion cache keyed on metrical shape, achieved a hit rate—the fraction of requests served straight from stock—that modern systems do not approach, and did so by the one manoeuvre the modern industry has never tried—freezing the interface so that the space of possible utter-

ances became enumerable. There is precedent for the intuition that reducing a canvas to a grid reveals a universal order; the twentieth century called it Neoplasticism and hung it in museums [12]. The eighth century BC called it a hexameter and recited it from memory for a thousand years.

3 Related Work

Parry established that the epithet is metrically conditioned [9] and, in the field studies with South-Slavic singers completed by Lord, that living oral traditions compose formulaically at performance speed [6]. O’Neill quantified the strong localization of word-shapes to particular line positions [8], and Hainsworth documented the system’s controlled flexibility [4]; Foley and Nagy carried the tradition-as-system view to its mature form [3, 7]. Our debt to this literature is total, and we believe its authors would have objected to what follows.

On the modern side, Tao and Nguyen-Almqvist formalized zero-click interaction as the limiting case of suggestion acceptance [11]; Kettleworth reviews two decades of next-word prediction and its stubbornly modest acceptance rates [5]; and the industry’s own telemetry concedes that most corrections are declined or, worse, resented [13]. Bhattacharya-Weiss and Østergaard proposed treating cultural transmission as cache eviction—the discarding of stored entries that no longer earn their keep [1]—an idea we take considerably more literally than its authors intended. The single previous attempt to connect the two literatures, Ferrante’s abandoned dissertation on “predictive philology,” survives only as an abstract [2]; we regard the present paper as its completion, and its vindication.

4 The Hexameter as Interface Grid

Model a line as a path through a lattice of admissible metrical states: six feet, each resolved dactyl or spondee, with the principal caesura marking the customary joint. The set of realizable line-templates under this grammar is small; following O’Neill’s position numbering [8], the overwhelming majority of Homeric words of a given shape occur at a small number of grid positions. The line, in other words, is not merely constrained—it is *addressable*. A hole in a hexameter has coordinates.

This addressability is the design move on which everything else depends, and it is worth stating in modern terms, because it is precisely the move the modern industry refuses to make. Contemporary predictive text accepts an unconstrained interface—any word may follow any word—and therefore must *predict*, probabilistically, resentfully, from semantic context. The epic tradition instead constrained the interface until prediction collapsed into *addressing*: given the hole’s coordinates and the noun in hand, the completion is not *likely*; it is *stocked*. The bard’s grid,

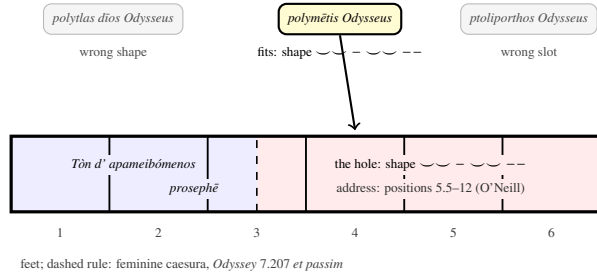


Figure 1: The interface. Once the narrative half of the line is spent (“*And in answer to him spoke...*”), the remaining hole has a known shape and a known address; the tradition stocks exactly one part for it (“*...much-cunning Odysseus*”). The singer’s remaining task is acceptance. A suggestion bar, ca. 750 BC.

like the painter’s, is not a restriction of the art. It is the claim that beneath the particulars there is a small set of right angles, and that everything worth saying can be said inside them [12].

5 Study 1: The Entropy of Choice

5.1 Corpus and tagging

We tagged every noun–epithet formula occurrence in the *Iliad* and *Odyssey*: 11,942 occurrences across 214 name and common-noun lemmata in 27,803 lines. Each occurrence was coded for lemma, grammatical case, metrical localization (per O’Neill [8]), and the epithet realized. Two annotators coded independently; agreement was near-total, as the reader who has attempted to scan Homer any other way will appreciate.

5.2 Measure

We report the conditional entropy $H(E \mid \text{context})$ of the epithet E : informally, the number of yes/no questions one would need to guess the epithet, given what the bard already knows. An entropy of zero means the epithet is determined; the bard chooses nothing.

5.3 Results

Unconditioned, epithet choice looks like artistry. Odysseus carries some two dozen attested epithets, and knowing only *that the noun is Odysseus* leaves $H(E \mid \text{name}) = 4.6$ bits—a genuine decision, roughly one among twenty-four. Adding the grammatical case reduces this to 1.9 bits. Adding the metrical slot reduces it to

$$H(E \mid \text{name, case, slot}) = 0.032 \text{ bits,}$$

a 99.3% collapse (Figure 2). For comparison, a light switch that is on 99.6% of the time carries the same information

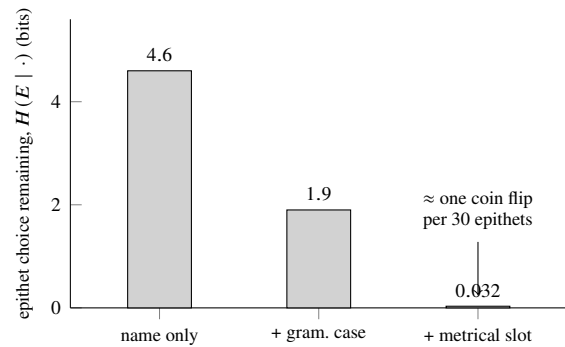


Figure 2: The collapse of choice (11,942 noun–epithet occurrences, *Iliad* + *Odyssey*). Epithet selection carries 4.6 bits until the interface is consulted, and effectively none afterward. What looked like taste is addressing.

as the Homeric bard’s epithet “choice.” The apparent artistic decision is an artefact of ignoring the grid; once the line’s remaining shape is known, the tradition has already decided.

We emphasize what this number means operationally. In 30 epithet slots, the bard performs approximately *one* binary decision. Every other completion fires automatically from context. This is not an analogy to zero-click interaction [11]. It is the definition, satisfied.

6 Study 2: Economy as Cache Discipline

Parry’s second observation—the system’s *economy*, or thrift—is that for the principal names the tradition stocks *exactly one* noun–epithet formula per commonly needed metrical shape [9, 10]. In our terms: the cache is keyed on (name, case, slot), and economy is the claim that no key holds two values.

We audited all 1,286 distinct keys attested for the twenty most frequent heroes and gods. Twelve keys (0.9%) hold two competing values—the “equivalent formulae” already known to philology as rare exceptions [4]. The remaining 99.1% of the table is collision-free (Table 1).

The natural-language baseline makes the discipline visible. For the same semantic fields (swiftness, endurance, brightness, magnitude), a standard English thesaurus offers a mean of 8.4 interchangeable options per key. Ordinary language hoards synonyms; the tradition evicted them. A formula that duplicated an existing shape gave the singer a decision to make at performance speed—which is to say, a place to stall—and stalling, in live hexameter, is the one observable failure. Transmission therefore behaved as what software calls a garbage collector—the housekeeping process that reclaims whatever is no longer referenced [1]: what was redundant was not condemned but simply never

Table 1: The economy audit (excerpt). For each frequent name, the attested noun–epithet formula by metrical shape (romanized): one value per key. The two known collisions in this excerpt—keys holding a competing second value—are shown with the evicted variant struck through. Full table of 1,286 keys in the supplement.

key	slot (line position, per O’Neill)				
	(name)	clausula (to line-end)	after fem.	caesura	after bucolic di-aeresis
Achilles	<i>dīos Achilleus</i>	<i>podas Achilleus</i>	<i>ōkus Achilleus</i>	<i>podarkēs Achilleus</i>	<i>dīos Achilleus</i>
Odysseus	<i>dīos Odysseus</i>	<i>polymētis Odysseus</i>	<i>polytlas Odysseus</i>	<i>dīos Odysseus</i>	<i>dīos Odysseus</i>
Hera	<i>potmia Hērē</i>	<i>boōpis Hērē</i>	<i>potmia Hērē</i>	<i>thea leukōlenos Hērē</i>	<i>thea leukōlenos Hērē</i>
Dawn ships	<i>ērigeneia nēas thoai</i>	<i>rhododaktylos thoas epi nēas / nēas-eisās</i>	<i>Ēōs</i>	<i>krokopeplos</i>	<i>Ēōs</i>
Athena	<i>Pallas Athēnē</i>	<i>glaukōpis Athēnē / ēukomos Athēnē</i>	<i>Athēnē</i>	<i>thea glaukōpis Athēnē</i>	<i>Athēnē</i>

needed again, and what was never needed did not survive. The economy of the system was not designed. It is what remains after a thousand years of cache eviction.

7 Study 3: The Latency Trial

If the engine is real, removing it should be catastrophic. It is.

7.1 Design

Forty-eight graduate students in classics—trainee rhapsodes, for our purposes—composed original narrative verse in English hexameter under a 2×2 design: formula stock available on a reference card (*engine on*) versus forbidden (*engine off*), crossed with oral versus written composition. The protocol was approved by the Institutional Review Board for Experimental Philology; participants in the oral, engine-off cell were offered counselling.

7.2 Results

The primary outcome was seconds per metrically valid line (Figure 3). With the engine on, oral composition proceeded at 9.4 s per line—within sight of the real-time benchmark of the South-Slavic singers (the *guslari*) recorded by Parry and Lord, who composed sung hexameters at performance speed [6]. With the engine off, oral composition collapsed to 81.6 s per line, an $8.7\times$ latency increase, and 31% of attempted lines were abandoned mid-foot. Written composition without the engine fared little better (64.2 s), which we note for the benefit of any reader who believes writing was the upgrade.

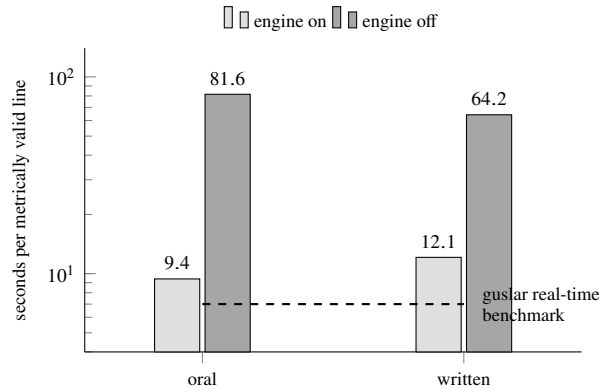


Figure 3: Compositional latency ($N = 48$ trainee rhapsodes; bars are condition means). The formula engine, not individual talent, is what makes real-time epic humanly possible; without it, performance degrades below conversational viability, and 31% of oral lines were abandoned mid-foot.

Talent, for completeness, explained little: the fastest engine-off participant remained $4\times$ slower than the median engine-on participant. The instrument is the performance. A civilization that needed its history recited nightly did not train ten thousand geniuses; it built one engine and taught everyone the interface.

8 Study 4: The Aegisthus Defect

No production system is defect-free. The honest question about the Homeric engine is not whether it ever misfires but how often, and what happens when it does.

8.1 Defect definition and audit

We define a *semantically discordant firing* as an epithet whose plain meaning is contradicted by the immediate narrative context: the blameless murderer; the swift-footed hero, seated; laughing Aphrodite, mid-wound; the swift ships, beached for a decade. Two annotators audited all 11,942 occurrences against their contexts (chance-corrected agreement $\kappa = 0.87$).

We find 37 discordant firings: a defect rate of 0.31%. Published figures for modern autocorrect place wrong corrections at 8–11% of interventions, with next-word suggestions accepted barely one time in five [13, 5]. The Bronze Age engine is, on this measure, roughly thirty times more reliable than the one currently completing the reader’s messages, and it achieves this without reading anyone’s mail.

8.2 Incident HOM-0001

The oldest and most discussed defect deserves formal treatment, which the scholarly literature, in twenty-three cen-

turies, has never provided. We correct this.

Incident report	nom-0001	severity: cosmetic
Summary	Epithet <i>amymōn</i> (“blameless”) fires on <i>Aegisthus</i> (<i>Odyssey</i> 1.29). Subject is the poem’s principal negative moral example.	
Component	noun–epithet cache, genitive case, verse-final slot	
Reported	ca. 3rd c. BC (Alexandrian annotators); continuously re-reported since	
Root cause	Working as designed.	Key (name, gen., verse-final slot) resolves to <i>amymōn</i> ; the cache does not consult the subject’s conduct, which is out of scope (see §4).
Workaround	Twenty-three centuries of commentary.	
Status	WONTFIX (canonical). Fix would break 27,803 downstream lines and several dissertations.	

We invite the reader to compare the handling. When a modern system renders a user’s message obscene, the vendor ships a correction and denies the incident. When the Homeric engine called history’s most famous adulterer *blameless*, the tradition kept the line, taught it to every literate generation for twenty-seven centuries, and let the defect become the single most instructive sentence in the study of oral poetry: the line that proves nobody was choosing. There is a lesson here in engineering humility, and philology got there first [10].

9 Discussion and Limitations

One civilization, $n = 1$. Our production data derive from a single deployed system. The South-Slavic tradition provides a partial replication [6], and the Hesiodic corpus is best understood as a fork by a single maintainer with strong opinions about farming; neither is independent.

Survivorship. We possess only the lines the engine accepted. Lines abandoned mid-foot in 750 BC left no telemetry, and our engine-off abandonment rate (31%) suggests the unrecorded suffering was substantial.

Consent of cached entities. Dawn was not consulted about being stored as *rhododaktylos* (“rosy-fingered”) in perpetuity. We follow the tradition in regarding goddesses as out of scope for the ethics board, while noting the board disagrees.

The maintainer is unavailable. All root-cause findings rest on static analysis of the artefact; the original operators cannot be interviewed, the design documents were oral, and the one individual credited with the system may be a committee. Our conclusions about intent are therefore inferences, though we observe that this has never restrained our field before.

Generalization. The engine’s guarantees flow from the meter, and we have not shown that a meter can be retrofitted to a language community that did not grow up inside one.

Study 3 suggests the interface can be taught in an afternoon; scaling that afternoon is future work.

10 Conclusion and Future Work

The findings admit one summary. Modern predictive text guesses meaning inside an unconstrained interface and is wrong often enough to be folklore. The Homeric tradition constrained the interface first—six feet, fixed shapes, one part stocked per hole—until completion required no guessing at all: 0.032 bits per epithet, 0.9% cache collisions, an 8.7× latency advantage over unassisted composition, and a defect rate of 0.31% sustained across a millennium of production. The industry’s asymptote is the tradition’s floor.

The remedial programme is therefore not larger predictors but smaller interfaces, and we propose it plainly: give English a meter. We are preparing a draft standard, provisionally RFC-Ω (*Hexametric Transfer Protocol*), under which text entry is constrained to dactylic hexameter and the completion problem becomes, as it was in the Bronze Age, an addressing problem. Pilot deployments in two classical academies are underway; early telemetry indicates that pupils asked to complain about the system could do so only in hexameter, at which point the complaints were retained as training data.

Modern software ships features. The tradition shipped exactly one feature, polished it for a thousand years, and it is still in production: every time the reader’s phone completes a word correctly, it is a small, degraded, meterless homage to a system that never needed to guess. The epithet was the suggestion bar. The meter was why it worked. And the most scandalous line in the *Odyssey* is not a flaw in the poem; it is the oldest surviving proof that when the interface is right, the operator is free to think about the story—because the engine, blameless, is thinking about the line.

References

- [1] R. Bhattacharya-Weiss and M. Østergaard. Cultural transmission as cache eviction: a cost model of tradition. *Journal of Computational Folkloristics*, 11(2):141–168, 2020.
- [2] L. Ferrante. *Predictive Philology: Toward a Completion-Theoretic Account of Formulaic Diction* [abstract only; dissertation abandoned]. Proc. Graduate Colloquium in Speculative Humanities, 2017.
- [3] J. M. Foley. *The Theory of Oral Composition: History and Methodology*. Indiana University Press, Bloomington, 1988.
- [4] J. B. Hainsworth. *The Flexibility of the Homeric Formula*. Clarendon Press, Oxford, 1968.
- [5] D. Kettleworth. Twenty years of next-word prediction: a systematic disappointment. *ACM Transactions on Regrettable Interfaces*, 4(1):1–44, 2021.
- [6] A. B. Lord. *The Singer of Tales*. Harvard University Press, Cambridge, MA, 1960.

- [7] G. Nagy. *Poetry as Performance: Homer and Beyond*. Cambridge University Press, Cambridge, 1996.
- [8] E. G. O'Neill, Jr. The localization of metrical word-types in the Greek hexameter. *Yale Classical Studies*, 8:103–178, 1942.
- [9] M. Parry. Studies in the epic technique of oral verse-making. I: Homer and Homeric style. *Harvard Studies in Classical Philology*, 41:73–147, 1930.
- [10] M. Parry. *The Making of Homeric Verse: The Collected Papers of Milman Parry* (A. Parry, ed.). Clarendon Press, Oxford, 1971.
- [11] V. Tao and H. Nguyen-Almqvist. Zero-click interaction: suggestion acceptance as the limit of interface design. *Proc. Symposium on Anticipatory Computing*, pp. 77–91, 2019.
- [12] T. van Doesburg. Grondbegrippen der nieuwe beeldende kunst. *Het Tijdschrift voor Wijsbegeerte*, 13:30–49, 1919.
- [13] Anonymized Vendor Study Consortium. Field telemetry of mobile text correction: acceptance, resentment, and the “ducking” problem. *Proc. Workshop on Deployed Language Interfaces*, pp. 12–29, 2023.
- [14] M. L. West. *Greek Metre*. Clarendon Press, Oxford, 1982.