

Intertextuality in the Digital Age¹

Neil Coffee, J.-P. Koenig, Shakthi Poornima, Roelant Ossewaarde, Christopher Forstall, and Sarah Jacobson

University at Buffalo, SUNY

When does Vergil refer to the work of another author, and what does it mean? Variations of this question have been asked about a range classical works from the height of Alexandrian scholarship in the 3rd century BCE to our own day. Formulations have changed, from ancient notions of imitation, emulation, and theft, to the modern emphasis on artistic repurposing and concerns with the ontological status of textual connections.² In classical studies, Latin poetry has been the subject of the most intensive investigation, but scholars of Greek poetry and of Greek and Latin prose have become increasingly interested in how authors reuse and refer to their predecessors and contemporaries.³

Digital approaches now in development can accelerate this research and provide new large-scale perspectives on intertextuality.⁴ These approaches are drawn from the discourse of digital humanities, and in particular from digital textual analysis. Humanities scholars have

¹ We would like to thank the graduate student participants in the fall 2010 seminar on intertextuality at the University at Buffalo, Kevin Roth, Sarah Skelley, and Valerie Spiller, for their contributions to the testing process described below. We would also like to thank Neil Bernstein, Stephen Hinds, Damien Nelis, Paul Roche, and the anonymous reviewers of *TAPA* for suggestions that greatly improved this piece.

² D'Ippolito 2000 reviews ancient conceptions of text reuse. Giangrande (e.g. Giangrande 1967) and his students have taken a lexical approach rooted in Alexandrian traditions that "as a rule, [does] not treat allusion as would a literary critic" (Farrell 1991: 14). Pasquali 1968 provided an influential exposition of allusion as artistic practice and is taken as a starting point by major works on Latin intertextuality such as Barchiesi 1984: 92, Conte 1986: 22-39, Thomas 1986: 171, Farrell 1991: 13, Wills 1996: 15 n. 2, and Edmunds 2001: 12. Hinds 1998: 20 n. 7 gives a précis of Pasquali's Anglophone reception. Conte 1986: 24-25 notes that that the German scholars under whom Pasquali studied were already engaged with the expressive possibility of intertextuality. Important intertextual work was also done in the early modern period. So, e.g., De La Cerda 1617 became the foundation for later work on Vergilian intertextuality, such as Knauer 1964.

³ Some examples are the collection Schepens and Bollansée 2005 on Polybius's use of his predecessors, Telo 2010 on Aristophanic intertextuality, Swift 2010 on the use of lyric features in classical Greek tragedy, and Levene 2010 on Livy's reuse of passages from his predecessors.

⁴ We use the word "intertextuality" to denote the broadest range of relationships between texts. For an account of the differences in terminology ("allusion," "intertextuality," "reference") used by scholars of Latin literature, see Farrell 2005, esp. p. 98 n. 2. We discuss definitions further in our concluding section.

increasingly used computational resources to engage in what Franco Moretti has called “distant reading,” or the use of quantitative models to examine trends in textual corpora.⁵ Distant reading of intertextuality has focused on identifying phrases or letter sequences in multiple locations that are exactly identical, or nearly so. Such matching techniques have been used to demonstrate that classical Roman and late antique authors favored different parts of Plato’s *Timaeus*, and that Diderot and d’Alembert borrowed heavily from the very Jesuits they were contesting in composing their 18th-century *Encyclopédie*.⁶ The study of intertextuality also encompasses a variety of textual relationships considerably more elusive than quotation. To address these with digital means requires identification of other computationally tractable features, in addition to exact word form or letter-sequence identity, that constitute intertextual phenomena. Bamman and Crane 2008 have demonstrated how detection of three features—word similarity (exact form and dictionary headword), word order, and syntactic similarity—could identify a range of potential intertexts, with the further suggestion of accounting for meaning, meter, and sound.⁷

This article reports on a comparison of one digital approach to intertextuality with traditional methods, and explores the consequences of automatic intertextual detection for practical research and theoretical horizons. Our comparison was performed using a free online intertextual comparison tool we created, available at <http://tesseract.caset.buffalo.edu/>. As tested, the Tesseract tool recovers approximately a third of the parallels captured by traditional

⁵ Moretti 2005: 1, who conducts literary historical inquiries that, *inter alia*, trace the development of novelistic sub-genres. Another form of textual analysis involves text mining to identify recurring themes that might escape the reader’s notice. A convenient example of the latter is the ongoing work of Cameron Blevins on the diary of Martha Ballard, a midwife living in Maine in the late 18th and early 19th centuries (<http://history.org/martha-ballards-diary/>).

⁶ Büchler, Geßner et al. 2010; Horton, Olsen et al. 2010.

⁷ Other digital efforts are afoot to further expand the range of criteria that can be captured. The eTRACES project, a collaborative effort funded by the German government through 2014, has the goal of producing advanced tools for intertextual study (<http://etraces.e-humanities.net/home-etraces.html>). The Italian Musisque Deoque project (<http://www.mqdq.it/mqdq/>) is endeavoring to build on its digital textual collection to produce a search engine for matching metrical patterns.

commentators, and adds a third not previously recorded. These results show that the tool can find valuable new parallels in even the best-studied authors, and suggest it could be particularly useful in identifying intertexts in works that have received less attention. In combination with information from commentators, our method allows us to draw conclusions about the overall intertextual artistry of ancient authors, as well as the intertextual reading habits of scholars.

In what follows, we begin by describing our search and testing processes. We then present a sample of our results showing how the Tesseract tool can be used to identify individual parallels and groups of parallels. From there, we explore the potential for distant reading opened up by automatic search, and conclude with observations on the consequences of digital approaches for intertextual study.

Search Method and Testing

Our search focused on parallels containing at least two similar words in each. This search was meant to approximate the traditional scholarly identification of *loci similes*, which takes two-word pairs as the most basic and common form of intertextuality. We left out of consideration the small and large-scale interactions that cannot be identified by lexical similarity alone. Vergil's *corruptique lacus, infecit pabula tabo* (*G.* 3.481)⁸ resembles Lucretius's *uastauitque uias, exhausit ciuibus urbem* (6.1140)⁹ in rhythm, sound, syntax, and general subject matter (plague), absent any shared words.¹⁰ These sorts of similarity were not captured by our search. We also did not address large-scale intertextual correspondences formed by thematic material or homologous structures such as book divisions.¹¹ Furthermore, our test considered

⁸ The plague “fouled the lakes and infected the pastures with disease.”

⁹ The plague “emptied the roads, and drained the city of its inhabitants.”

¹⁰ Thomas 1986: 180 = Thomas 1999: 122.

¹¹ Fantuzzi and Hunter 2004: 467 discuss Catullus 64 as an example of structural rather than lexical imitation.

only intertextuality between Latin texts, without accounting for the critical influence of Greek literature on Latin.¹² Rather than attempting to address the full complexity of intertextuality at once, we chose instead to explore what automatic detection of a key subset of intertextual phenomena could achieve, and whether it could constitute a step toward automatic detection and analysis of the full range of intertextual phenomena.

To evaluate our search process, we performed two tests comparing book 1 of Lucan's *Civil War (BC)* with all of Vergil's *Aeneid*. These works were chosen because they were long enough to give representative results, and well-studied enough to allow comparison with traditional approaches. Each of our two tests was conducted with a different algorithm.¹³ Version 1 search found minimum two-word phrases with exactly identical words in either order, separated by no more than four other words. So Lucan's *horrida . . . dumis* (1.28) describing Italian fields overgrown during civil war matched Vergil's *dumis . . . / horrida* (9.381-2) referring to obstacles to the flight of Nisus and Euryalus. Version 2 matched minimum two-word phrases by dictionary headwords (lemmas), in any order, with sentences as the unit of comparison.¹⁴ In the Version 2 search, Lucan's *notae fulsere aquilae* (1.244) referring to the glimmer of Roman eagles that frightens the citizens of Ariminum matched Vergil's *notis fulserunt cingula bullis* (12.942) describing the glimmer of the baldrick of Pallas that incites Aeneas to kill Turnus. Test results were compared to parallels compiled from commentaries to explore the efficacy of the digital intertextual search and inform site users of the kinds of results

¹² The necessary techniques for cross-linguistic matching are still in development; see the closing section of this piece. The Tesseract site currently offers Greek-to-Greek matching of a limited number of texts, and Latin-to-Latin matching of a larger number. Current plans foresee substantially increased availability of texts in both languages in the near future.

¹³ Both algorithms are available on the Tesseract website, <http://tesseract.caset.buffalo.edu>, as versions 1 and 2.

¹⁴ We used the Archimedes Project Morphology Service (<http://archimedes.mpiwg-berlin.mpg.de/arch/doc/xml-rpc.html>, accessed August 21, 2011) to identify the lemmas. The parser sometimes returned more than one lemma for a given word. In the case of multiple lemmata for one word, all lemmata were checked for matches, leading to some false positives. So, e.g., *bello* returned the adjective *bellus*, *a*, *um*, as well as the noun *bellum*, *i* as possible lemmata.

they could expect. The current third version available on the website should provide faster and more accurate search, but awaits further testing.

Results: Individual Parallels and Sets of Parallels

The last example above was found by our lemma search, but does not appear in any of the major commentaries on Lucan, including the ample new volume on *BC 1* by Roche.¹⁵ We start the discussion of our results with consideration of this example to demonstrate how computational methods can reveal new instances of the sort of local correspondences that have traditionally interested scholars even in well-studied texts.

If we look for sources for Lucan's *notae fulsere aquilae*, we find no phrase consisting of forms of the noun *notae* and the verb *fulgeo* in extant Latin before Vergil.¹⁶ When Lucan uses the phrase to describe Caesar's invasion of Italy, he starts his war narrative with a complex of emotions and perspectives taken from the very end of the *Aeneid*, where Aeneas kills Turnus. In general, Lucan's connection of his opening with the end of the *Aeneid* suggests that features of civil war found in the *Aeneid* were not ended with the close of Vergil's epic but rather perpetuated down to the republic. More specifically, Aeneas's recognition of the bosses (*notis . . . bullis*) suggests that his motivations for killing Turnus are allegiance and sympathy toward Pallas. In *BC 1*, the resplendence of arms highlights the tragically misdirected excellence of

¹⁵ Roche 2009. Other commentaries and articles consulted are: Heitland and Haskins 1887, Thompson and Bruère 1968, and Viansino 1995. As the length of commentaries continues to expand (e.g. Horsfall 2000 at over 567 pages), Roche looks slim by comparison, but he offers a wealth of parallels, the greatest number of any Lucan commentator by far (see further below).

¹⁶ Silius Italicus later employs the same phrase in the *Punica* to describe the arms of a Roman mistakenly killed by his son the night before the battle of Cannae: *notis fulsit lux tristis ab armis* (9.107) "The mournful light shone from the well-known arms."

Roman troops. Amid the institutional forces driving civil war, the personal ties that animate Aeneas have little influence.¹⁷

Because automatic detection reveals numerous parallels at once, it can be particularly useful in identifying constellations of intertexts. One such configuration involves a single passage in a later text containing multiple parallels with an earlier text. When Lucan's Caesar is about to make his momentous crossing of the Rubicon and invade Italy, he is suddenly brought up short.

ut uentum est parui Rubiconis ad undas,
ingens uisa duci patriae trepidantis imago
clara per obscuram uoltu maestissima noctem
(*BC* 1.185-7)

As Caesar approached the waters of the slender Rubicon,
Before the general appeared a giant specter of his shaken fatherland,
Shining brightly through the night's gloom, but downcast.¹⁸

Three commentators noticed that words uisa . . . maestissima recalled the words maestissimus Hector / uisus of *Aeneid* 2.270-1, where the ghost of Hector appearing in a dream to Aeneas corresponds to the specter of Roma appearing to Caesar at the Rubicon.¹⁹ Roche notes that, unlike Aeneas, Caesar is in the midst of creating the crisis of which the vision warns, and disregards the warning. The intertext thus contrasts Caesar's indifference to the destruction he wreaks with Aeneas's deep concern for Trojan losses.²⁰ Searches by word form and lemma revealed two further *Aeneid* parallels not noted by commentators that support this contrast. Lucan's phrase Patriae trepidantis imago (1.186) echoes the words patriae strinxit pietatis

¹⁷ Lucan's contradictions of Vergil in fact rework readings available in the *Aeneid*. The futility of Aeneas's allegiance with Latinus and Aeneas's own inability to save Pallas also suggest the feebleness of personal ties in the face of larger historical processes. Lucan's primary difference from Vergil on this theme is in his emphasis on impersonal social forces, in contrast to Vergil's invocation of divine causation and fate.

¹⁸ All translations are our own unless otherwise noted.

¹⁹ Thompson and Bruère 1968, Viansino 1995, and Roche 2009.

²⁰ Roche 2009 ad 1.186 observes in addition that the phrase uisa . . . imago is commonly used to introduce apparitions in epic, appearing several times in the *Aeneid*: 2.773, 4.557, 5.636-7, 5.722.

imago (*Aeneid* 9.294, also with the *imago* at line end) describing the filial affection that inspires Ascanius to promise to care for the mother of Euryalus as the young warrior sets forth on a dangerous mission. Vergil uses the same phrase to describe the pity that Aeneas feels for Lausus, whom Aeneas killed as he rushed to avenge his father: *mentem patriae subiit pietatis imago* (10.824). Lucan's *Patria* refers to Rome's tutelary goddess, while Vergil's term is in both instances an adjective, *patrius, a, um*, "fatherly." Whereas Caesar in the end simply disregards his apparition, Ascanius responds by committing a generous act and Aeneas is moved to sympathy.

Individually these parallels highlight Caesar's indifference; collectively they foreground his social isolation. The relations between Hector and Aeneas, Aeneas and Ascanius, and even Aeneas and Lausus, are deeply personal. Caesar's encounter with the goddess Roma points rather to his lack of human social ties anywhere in the poem: he is approached only by an abstract divinity, not by any friend, son, or even pitiable enemy. The closest Caesar comes to a personal relationship is his desire to possess Rome itself.

Constellations of intertexts can occur not just in one phrase, but across various locations in the receiving text. Thus we find four separate instances in *BC* 1 where Lucan recalls the *Aeneid* while referring to Italian city walls (*moenia*). The second example was found by our automatic Tesserae search and commentators, the other three only by automatic search.

1. Lucan's mention of tottering house walls in Italian cities at the opening of his epic (*semirutis pendent quod moenia tectis / urbibus Italiae*, 1.24-5) reverses the notion of progress through the foundation of Rome's walls at the opening of the *Aeneid* (*dum conderet urbem // . . . altae moenia Romae*, 1.5-7).

2. As Lucan's divine Roma appeals to Caesar to stop at the Rubicon, she invokes Capitoline Jupiter guarding Rome's walls (*[Iuppiter] magnae qui moenia prospicis urbis / Tarpeia de rupe*, 1.195), recalling the apparition of Hector encouraging Aeneas to seek new walls for his people (*his moenia quaere / magna*, 2.294-5). Lucan's reference here is concurrent rather than adversative: Caesar threatens the great and storied walls of long foundation.
3. When the people of Rimini, alarmed by the prospect of invading Roman soldiers, complain that their city is too close to the dangerous Gallic border (*o male uicinis haec moenia condita Gallis*, 1.248), their lament pointedly reverses Aeneas's admiration for the rising walls of Carthage (*o fortunate, quorum iam moenia surgunt!*, 1.437). The walls of Ariminum thus seem far from providing the sort of secure home Aeneas longs for. Lucan's *moenia condita* at 1.248 also evoke the *telos* of founding Rome announced at the opening of the *Aeneid* (*dum conderet urbem*, 1.5; *altae moenia Romae*, 1.7), as do the other passages here.
4. Caesar riles up his soldiers by suggesting they have been denied land due for military service (*quae noster ueteranus aret, quae moenia fessis?*, 1.345). His phrasing recalls the words of Aeneas's senior counselor Nautes, who advises allowing Trojans exhausted from wandering to settle in Sicily (*et his habeant terris sine moenia fessi*, 5.717). The concern of Nautes and Aeneas for their people contrasts with Caesar's use of one argument among many to advance his personal interests. Roche also notes the parallel between this phrase of Lucan and *Aeneid* 3.85, where Aeneas implores the god Apollo to direct him to the new Trojan homeland with the words *da moenia fessis*.

The clear verbal similarities of the last parallel make it the most individually compelling, but all four together create a small program of contrast between *BC 1* and the *Aeneid*. Vergil's *moenia*

suggest permanence, solidity, and safety, the achievement of a home and a refuge. Lucan writes of the collapse of city walls, their failure to keep citizens safe, and their utility only in rhetoric.²¹ Lucan's repeated contrast undermines Vergil's image of the city as a secure refuge and suggests Aeneas's efforts to found one are pointless.²² The complex of allusion prepares the way for Lucan's extended development of the paradox that, at Caesar's approach, Romans seek safety by abandoning their city walls rather using them for defense (1.484-520).

Large-scale Intertextuality

We hope to have demonstrated that automatic search can reveal the sort of individual intertexts and groups of intertexts typically found to be meaningful. An evaluation of our results requires not just individual examples of successful searches, however, but a full-scale comparison with parallels found by commentators.

Amid the many studies addressing classical intertextuality, few have examined these relationships in the aggregate. So Hinds:

Modern scholarship of allusive relationships can be broadly divided . . . into studies of local contact (which tend to bracket out more systematic implications) and studies of systematic contact (which tend to bracket out details of local contact). In line with the Latinist's traditional preference for concrete and isolable effects over intertextual open-endedness of any kind . . . , purely local approaches have usually been predominant.²³

In the field of classics, the major exceptions are Knauer's study of the *Aeneid* and Homer and Nelis's study of the *Aeneid* and Apollonius.²⁴ Each of these works contains a wealth of individual analyses and summative judgments, but studies of this type are rare precisely because

²¹ Walls are a theme of *BC* 6 as well, as discussed by Saylor 1978.

²² Again, Vergil anticipates Lucan. The fact that all the *Aeneid* passages in this constellation are from the poem's first half is instructive. By the time that Lavinium is burned in *Aeneid* 11, the ideal of a secure city has receded.

²³ Hinds 1998: 101.

²⁴ Knauer 1964 and Nelis 2001. Casali 2011: 81 n. 2 writes that "a study on the intertextual relationship between Lucan and Virgil analogous to Nelis (2001) would be desirable." We hope that the work presented here offers a step in that direction, though along a different path.

few scholars have the requisite skills, industry, and endurance to undertake such vast projects. Approaching these studies can also be sufficiently daunting that even scholars explore selectively, and so still obtain only a partial picture of full intertextual relationships. To comprehend multiple such volumes is an overwhelming prospect for even the most capable critics.²⁵

One way out of this aporia is to develop methods of describing intertextual patterns that can be absorbed at once, and that can be played off against readings of individual *loci*. That is, we can employ forms of distant reading to complement traditional close reading. In Moretti's conception of distant reading,

distance is . . . not an obstacle, but a specific form of knowledge: fewer elements, hence a sharper sense of their overall interconnection. Shapes, relations, structures. Forms. Models.²⁶

Engagement with distant reading may require an adjustment in perspective for literary scholars accustomed to treating each individual instance with rigorous scrutiny. Moretti's methods assume a level of comfort with the quantitative patterns of inference native to the natural and social sciences that allow for abstraction of trends from large sets of individual instances. But we can hope that by working back and forth from distant to close reading, it will be possible for scholars to arrive at satisfying synthetic interpretations of intertextuality without being overwhelmed by its totality. No less importantly, simplified representations of intertextuality can help make intelligible to students what would otherwise seem an arcane element of language and literary artistry.

²⁵ Farrell 2005: 107: "Of course the mind recoils from the thought of a library full of books entitled, 'The *Aeneid* and Homer,' 'The *Aeneid* and Apollonius,' 'The *Aeneid* and Ennius,' and so forth. At the same time, however, I at least have no doubt that many such books could be written."

²⁶ Moretti 2005: 1, emphasis his.

Automatic Detection: Ranking System

Our own development of distant intertextual reading began from two basic questions. First, can our method detect a significant number of meaningful parallels? Second, if we were assured that we had a substantial body of meaning parallels, what observations could we make about them? To answer the first question, we manually scored our automatically generated results along with those from a set of commentators. Scoring both sets together helped to ensure that we applied a consistent standard to each. We devised the following system for our ranking.

Tesserae Intertextual Parallel Scoring System

| Meaningful | | Not Meaningful | | |
|--|--|---|--|---|
| Interpretable | | Not Interpretable | | |
| More significant | Less significant | | | |
| 5 | 4 | 3 | 2 | 1 |
| <ul style="list-style-type: none"> • High formal similarity in analogous context. | <ul style="list-style-type: none"> • Moderate formal similarity in analogous context, or • High formal similarity in moderately analogous context. | <ul style="list-style-type: none"> • High / moderate formal similarity with very common phrase or words, or • High / moderate formal similarity with no analogous context, or • Moderate formal similarity with moderate / highly analogous context. | <ul style="list-style-type: none"> • Very common words in very common phrase, or • Words too distant to form a phrase. | <ul style="list-style-type: none"> • Error in discovery algorithm, words should not have matched. E.g. phrase across punctuation boundary judged not meaningful. |

“Formal similarity” refers to any linguistic features of the words identified or the phrases they constituted, including those searched for (word and stem similarity) and others not specifically searched for (e.g. meter, sound, syntax).²⁷

²⁷ For his large-scale study of the relationship between the *Aeneid* and Homer, Knauer 1964 ranked parallels based upon his judgment of similarity (translated into symbols for different types of parallels on p. 364). We attempted to

The main purpose of the scoring system was to distinguish between meaningful (5-3) and not meaningful (2-1) parallels. This distinction follows that of Thomas, who writes that intertexts are either “susceptible to interpretation or meaningful” or not.²⁸ The first group ranked not meaningful included those parallels that did not meet our intended search criteria due to limitations of our search process, which were given a score of 1. For example, false positive matches appeared when a given form in the text had more than one possible lemma. Thus *BC* 1.52 *uelis . . . regnum* matched *Aeneid* 4.587-91 *uelis /// . . . regnis*. Lucan’s *uelis*, referring to Nero’s preference of position among the stars, is a form of the verb *uolo*, whereas Vergil’s noun *uelis* refers to the sails of Aeneas departing from Carthage.

Another group, those ranked with a 2, were not returned erroneously, but were nevertheless judged not meaningful because the two words of the phrase were so far apart or so ordinary that we believe few critics would recognize them as a parallel or choose to investigate them further. So, for example, the following instance:

certe populi quos despicit Arctos
felices errore suo, quos ille timorum
maximus haut urguet leti metus.
(*BC* 1.458-60)

To be sure the pole star looks down upon peoples
fortunate in their error, for the greatest of fears,
the fear of death, oppresses them not at all.

Pandarus, ut fuso germanum corpore cernit
et quo sit fortuna loco, qui casus agat res,
portam ui multa conuerso cardine torquet
obnixus latis umeris.
(*Aeneid* 9.722-5)

Once Pandarus realized his brother was strewn on the ground
and recognized the adverse circumstances, what misfortune was upon them,

be more precise by adhering to the two criteria of formal similarity and context analogy, though of course these were subject to a range of interpretation as well.

²⁸ Thomas 1986: 117.

he turned the gate on rotated hinge with all his might, pushing against it with his broad shoulders.

As critics have reminded us,²⁹ whether a potential moment of intertextuality appears interpretable depends upon the subjectivity of the reader: can he or she see some similarity between the two loci that admits further significance? That may indeed be possible in instances such as this.³⁰ Researchers investigating linguistic patterns could potentially examine such parallels profitably, and indeed we hope to enable such linguistic study. Critics investigating literary style could evaluate the similarity of two instances of anaphora in the forms of *qui . . . qui*. But in the regular course of literary interpretation, positing any link between the passages based on such common words, commonly expressed, would, in the words of Hinds, put “strains . . . upon a philological decorum of interpretability.”³¹ Conversely, for a phrase to be declared meaningful and be given a score of at least 3, it needed to represent a coherent, distinctive idiom. This criterion was meant to exclude phrases that consisted of pairs that were so common as to be colorless (e.g. *aut . . . aut*) or so widely separated that they did not constitute or occur in one phrase.

Among the meaningful parallels discovered (types 5-3), we distinguished between passages that generated new significance through their association (types 5-4) and instances where text was reused without creating such significance (type 3). We have called the former “interpretable” and the latter “not interpretable,” following a distinction made by critics. In the case of Lucan, for example, Roche writes of the “incidental influence of the language and

²⁹ Notably Martindale 1993.

³⁰ Hinds 1998: 26: “There is no discursive element in a Roman poem, no matter how unremarkable itself, and no matter how frequently repeated in the tradition, that cannot in some imaginable circumstance mobilize a specific allusion. This is a truth often suppressed by professors of Latin for reasons of pedagogy and (perhaps) peace of mind; but it is a truth none the less.” Wills 1996 explores this subject in depth.

³¹ Hinds 1998: 45.

expression of Senecan prose upon Lucan's Latin."³² Roche's "incidental influence" indicates an instance of mere language reuse, without the generation of meaning by the interaction of the two loci. We attempted to make this distinction more precise by introducing surrounding context as a criterion. To be deemed potentially generative of meaning, and so interpretable (5-4), a given parallel needed to have analogous contexts in both works that were not wholly typed. This criterion likewise originates in current critical practice. In a study of intertextuality in Martial, Hinds asks:

When Martial's elegiac couplets speak the cadences of Ovidian elegy, when and how far are the echoes to be read as thematically grounded, when or how far as utterly indifferent to content or context?³³

Hinds's "thematically grounded" parallels are equivalent to those that we deem generative of meaning, and so interpretable, and those that are "indifferent to content or context" are equivalent to simple language reuse, and so not interpretable.

We can also describe the distinction between interpretable (5-4) and non-interpretable (3) parallels with respect to genre. In an influential formulation, Conte has distinguished between a source text as "exemplary model," where the referring author directs the reader's attention to a particular moment in another work, and as a "code model," where the referring author draws instead on a set of generic gestures laid out by another author or authors.³⁴ We would suggest that the parallels we have classified as interpretable (types 5-4) correspond roughly to Lucan's use of Vergil as an "exemplary model," in that they generate meaning by evoking specific loci in the *Aeneid*. Instances of Lucan's use of Vergil as a "code model" generally correspond to text reuse we have designated not interpretable (type 3), in that they most often consist of distinctive

³² Roche 2009: 29. Cf. his statement at p. 24 that "frequently [Lucan's] diction is influenced directly or indirectly by Horace's own expression."

³³ Hinds 2007: 119.

³⁴ Conte 1986: 31.

language repeated by Lucan, but in phrases that do not evoke any particular resonance. When this language comes only from the *Aeneid*, or when it is stamped by Vergilian influence, it can be considered an instance of Lucan using epic language for its generic qualities without evoking a meaningful analogy, or a use of epic “code.” Some of the parallels we have classified under text reuse may simply be common phrases, metrically compliant but otherwise not particular to epic. But some preliminary inquiries into this topic have suggested that most *BC* 1 phrases shared with the *Aeneid* are found first there.

The analysis of the following parallel, scored as a 3, shows how these criteria were applied.

patriae sedes remeamus in urbis,
 inopiaque in medio peraguntur bella senatu.
 (*BC* 1.690-1)

But now I return to the site of the city,
 where the insidious civil war is conducted in the senate itself.

haud procul hinc saxo incolitur fundata uetusto
urbis Agyllinae sedes, ubi Lydia quondam
 gens, bello praeclara, iugis insedit Etruscis.
 (*Aeneid* 8.478-80)

Not far from here lies the site of the city of Agyllina,
 founded upon ancient stone, where once the Lydian people,
 renowned warriors, settled in the Etruscan hills.

This parallel avoided being assigned a score of 2 because its two words were close enough in each text and infrequent enough to constitute an instance of text reuse. On the other hand, the words are rather ordinary, and their arrangement in the line differs substantially between the two passages. Finally, it is difficult to find analogous elements in the topic and context of the passages that would make this a meaningful recollection by Lucan. In Lucan’s passage, the bacchant who predicts civil war for Rome returns in her prophetic and visionary journey to foretell dissent in the Roman senate itself. In Vergil’s passage, Evander describes the city of

Agyllina, where Aeneas may find allies willing to fight against Mezentius and the Latins. Beyond a context of war shared by the epic poems as a whole, there seemed little obvious analogy, concurrent or adversative, between the passages.

The final and most subjective distinction we employed was between stronger (5) and weaker (4) forms of meaning-generating parallels, which we called respectively “more significant” and “less significant.” This was an effort to indicate, among all meaning-generating parallels, those that seemed most meaningful according to traditional critical standards. The example given above of the parallel between Lucan’s *notae fulsere aquilae* (1.244) and Vergil’s *notis fulserunt cingula bullis* (12.942) represents a strong parallel that merited a score of 5. The following parallel earned a rank of 4.

tum rura Nemetis
qui tenet et ripas Atyri, qua litore curuo
molliter admissum claudit Tarbellicus aequor,
signa mouet.
(*BC* 1.419-22)

Then the soldiers who guarded the Nemes changed camp, and those holding the banks of the Atyrus, where the Tarbellian people surround the sea that flows gently to their curving shore.

feror huc et litore curuo
moenia prima loco fatis ingressus iniquis
Aeneadasque meo nomen de nomine fingo.
(*Aeneid* 3.16-18)

I am borne here [to Thrace] and place my first walls
on the curving shore, destined for frustration,
and for my people fashion the name of Aeneadae from my name.

Apart from the exception noted in our discussion of the *moenia* theme, the passages presented above have all been found only by the Tesseract search. In this case, we give an example of one found by Roche as well, to illustrate why it falls short of the highest ranking. Roche notes only that we find the words *litore curuo* ending lines at *Aeneid* 3.16, as well as *Aeneid* 3.223, Ovid

Fasti 3.469, and Valerius Flaccus 1.275.³⁵ His lack of fuller exposition suggests that he finds little substantial to explicate between the two passages, which might, in the terms used here, earn the echo the ranking of a 3 for pure language reuse. In our reading, it deserved a higher score both because of the exact repetition of the phrase, and because of a potentially analogous context. The analogy is between Aeneas founding a fortification he will have to abandon, and Roman legionaries in *BC* 1 abandoning their fortifications for civil war, both on or near curved shores. The parallel fell short of a ranking of 5, however, because the connection between the words *litore curuo* and this analogous context seemed slight.

In light of continued critical reminders of the subjectivity of meaning-making by the reader, the classification of parallels by the elementary criteria we have outlined may seem crude. Some classification system is required, however, to make a basic distinction between more and less useful results; to evaluate results from automatic discovery against those from traditional methods; and to present the outcome in intelligible form. More subtle distinctions—including those we posit between pure language reuse and different types of meaning-generating parallels—will be open to debate, but we hope our criteria will help bring greater transparency to the major types of intertextual phenomena. Our results will be posted online, and we hope to develop a system to allow users to comment on marked-up texts, so that future classification can benefit from the insights of a broader interpretive community.

Results of Automatic Detection

Applying our ranking system to Tesseract results and those of the commentators produced the following outcome.

³⁵ Roche 2009: 288.

Intertextual Parallels Between Lucan *BC* 1 and *Aeneid* Found by Tesseractae and Commentators, by Type
Testing Conducted Fall 2009 – Fall 2010

| TYPE | ALL | Tesseractae Word & Lemma | Tesseractae Word | Tesseractae Lemma | All Commentators | Heitland & Haskins | Thompson and Bruère | Viansino | Roche |
|-------|------|-----------------------------|---------------------|----------------------|---------------------|-----------------------|------------------------|----------|-------|
| 5 | 103 | 36 | 18 | 26 | 93 | 4 | 14 | 30 | 85 |
| 4 | 115 | 57 | 25 | 43 | 79 | 3 | 12 | 18 | 66 |
| 3 | 425 | 280 | 41 | 262 | 192 | 6 | 13 | 33 | 168 |
| 2 | 2289 | 2241 | 70 | 2182 | 55 | 1 | 1 | 8 | 50 |
| 1 | 486 | 486 | 5 | 481 | 0 | 0 | 0 | 0 | 0 |
| TOTAL | 3418 | 3100 | 159 | 2994 | 419 | 14 | 40 | 89 | 369 |
| 5-3 | 643 | 373 | 84 | 331 | 364 | 13 | 39 | 81 | 319 |
| 5-4 | 218 | 93 | 43 | 69 | 172 | 7 | 26 | 48 | 151 |

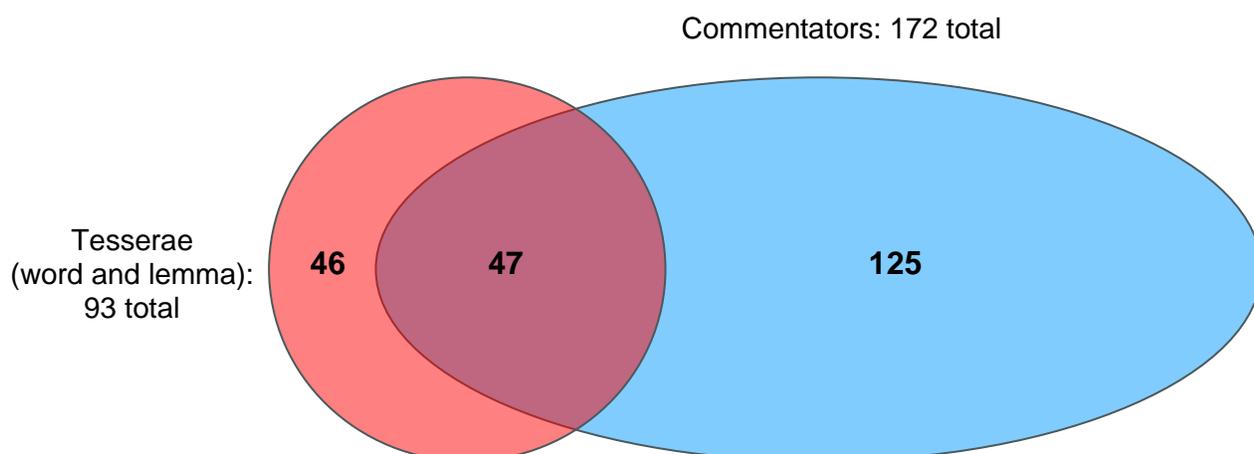
ALL Total of all Tesseractae and commentators
 Tesseractae Word Tesseractae identical word form match
 Tesseractae Lemma Tesseractae lemma match
 All Commentators Total of all following commentators
 Heitland and Haskings Heitland and Haskins 1887
 Thompson and Bruère Thompson and Bruère 1968
 Viansino Viansino 1995
 Roche Roche 2009

Note: in totals numbers, individual parallels found by more than one source are counted only once, so totals may be smaller than the sums of the individual values.

The table indicates that automatic discovery with manual examination of results can reveal substantial numbers of the parallels most significant to literary interpretation. The number of such parallels discovered naturally increases when we search for similar lemmas in whole sentences rather than exact words search in smaller windows, though the large number of less meaningful parallels returned by lemma search significantly increases the labor of sorting. If we compare the numbers of the most valuable parallels with the commentators, we find total numbers for types 5 and 4 from word search (43) and lemma search (69) comparable to those of Viansino (48), but fewer than half those of Roche (151). If we combine Tesseract word and lemma counts and exclude duplicates, the total (93) comes somewhat closer to Roche. Tesseract returns fewer valuable results because it does not take into account a variety of features that commentators account for when identifying parallels, such as semantics and syntax. We expect the Tesseract yield to increase as new search criteria are added.

For the moment, the automatic search available on the Tesseract website can serve as a check on and complement to traditional methods. More significantly, it can identify parallels unrecognized by commentators in substantial numbers, as illustrated in the following chart.

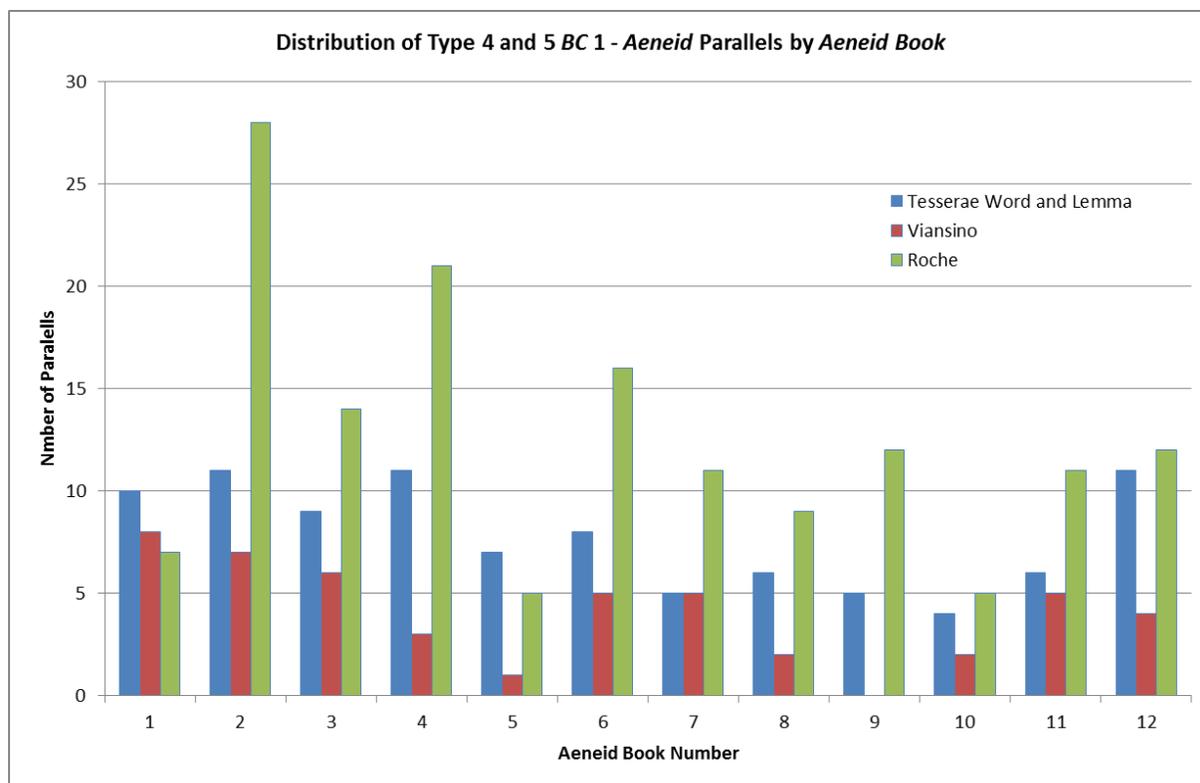
Unique Type 5 and 4 *BC 1* – *Aeneid* Parallels Found by Tesserae vs. All Commentators



The circle on the left represents all type 5 and 4 parallels (93) found by both Tesserae search methods (exact word and lemma). The oval to the right represents the total of the same types found by all commentators combined (172). The numbers within show, from left to right, parallels found only by Tesserae (46), by both Tesserae and commentators (47), and only by commentators (125). Although the Tesserae search captured only a quarter of the high interest type 5-4 parallels found by the commentators, this recovery rate nevertheless shows the capacity of automatic search to replicate traditional methods. Equally compelling is the fact that Tesserae search returned a high proportion of previously undiscovered parallels. This result suggests that the systematic application of fixed criteria can detect parallels traditional criticism may tend to overlook.³⁶

We can explore this last question in greater depth by considering which parts of a source text are referred to by the target text, or, in the case of our sample, how often intertexts from each book of the *Aeneid* appear in *BC 1*.

³⁶ A point made by Bamman and Crane 2008.



The chart above gives the number of interpretable intertexts between *BC 1* and individual books of the *Aeneid*, as found by each of three sources: combined Tesseract word and lemma search, Viansino, and Roche. The y-axis gives the number of parallels found, while the x-axis shows which book of the *Aeneid* the parallel occurs in. Each source finds more parallels to the first six books of the *Aeneid* than to the last six. Roche finds more parallels overall. The distribution of Roche's parallels in the *Aeneid* varies more than those found by Tesseract and Viansino, as indicated by the large gap between the highest and lowest numbers of parallels Roche finds (28 to *Aeneid* book 2 vs. 5 each to *Aeneid* books 5 and 10).³⁷

Where all sources find consistent variation in intertextual connections with different books of the *Aeneid*, this likely represents a real difference in Lucan's practice. Just as Servius devotes a much greater portion of his *Aeneid* commentary to the first half of Vergil's epic than

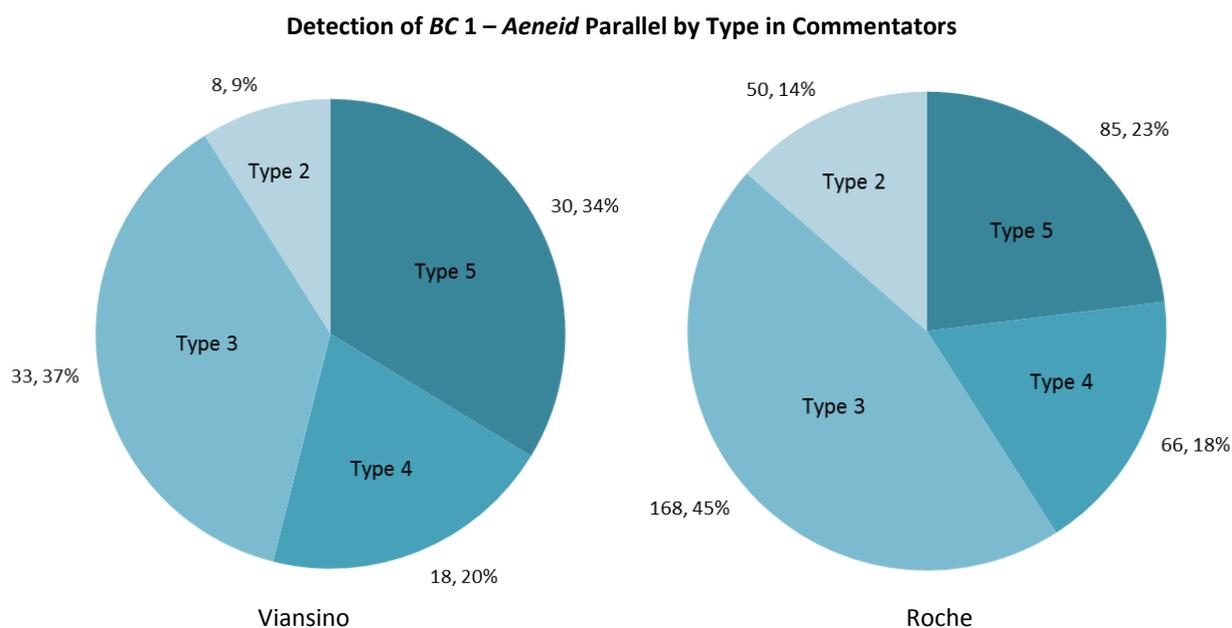
³⁷ By the ordinary measure of standard deviation, Roche's parallels fluctuate three times as much as those of Tesseract or Viansino. Roche has a mean number of high value parallels per *Aeneid* book of 12.6, with a standard deviation of 6.7. For Tesseract the values are respectively 7.8 and 2.6, and Viansino 4 and 2.4.

the second, Lucan may have had greater interest in, or use for, the first half of the *Aeneid* in composing *BC* 1. Both may have responded in a similar way to the work itself or been influenced by an inherited interpretive emphasis on *Aeneid* 1-6. We explore these possibilities further below when we combine results from all sources to obtain an overall picture of Lucan's intertextual patterns.

Where the sources diverge, comparison can illuminate their different critical practices. The principal difference is the greater variability in the results of Roche, who finds a high proportion of parallels between *BC* 1 and the perennially popular and well-studied *Aeneid* books 2, 4, and 6. Roche's higher variability may be a matter of outlook and interests—he had certain books of the *Aeneid* uppermost in his mind, while Viansino had others. Alternatively, the difference in variability may be a function of different methods. Roche conducts a deep analysis of *BC* 1, while Viansino and Tesseræ find fewer results, Viansino because he treats all of *BC*, and Tesseræ because it did not capture all commentator intertexts. So a deeper look at *BC* 1 may simply reveal greater variation in Lucan's intertextual practice. The development of more comprehensive automatic search system may help distinguish more clearly between observer preferences and textual differences.

If we consider not just interpretable parallels (5-4), but all meaningful parallels (adding type 3), we find another significant trend. More recent investigations, both by traditional commentators and our automatic search, have returned a significantly higher proportion of pure language reuse / code model language (type 3) intertexts, as well as ordinary language (type 2). These types typically represent, respectively, literary craftsmanship at work and prevailing linguistic patterns. The increase in type 3 and type 2 parallels would seem to be a natural result of subsequent commentators examining *BC* 1 more minutely. It would also seem to result from

the increasing use of digital search. In the latter respect, Roche represents something of a mid-point between earlier commentators and Tesseractae. The promotional materials for Roche's commentary point out that his was the first commentary on *BC 1* to make full use of electronic search.³⁸ The consequences are evident from comparison with Viansino. As the following charts show, more than half of Viansino's parallels are interpretable (types 5 and 4), whereas types 2 and 3 make up more than half of Roche's instances.

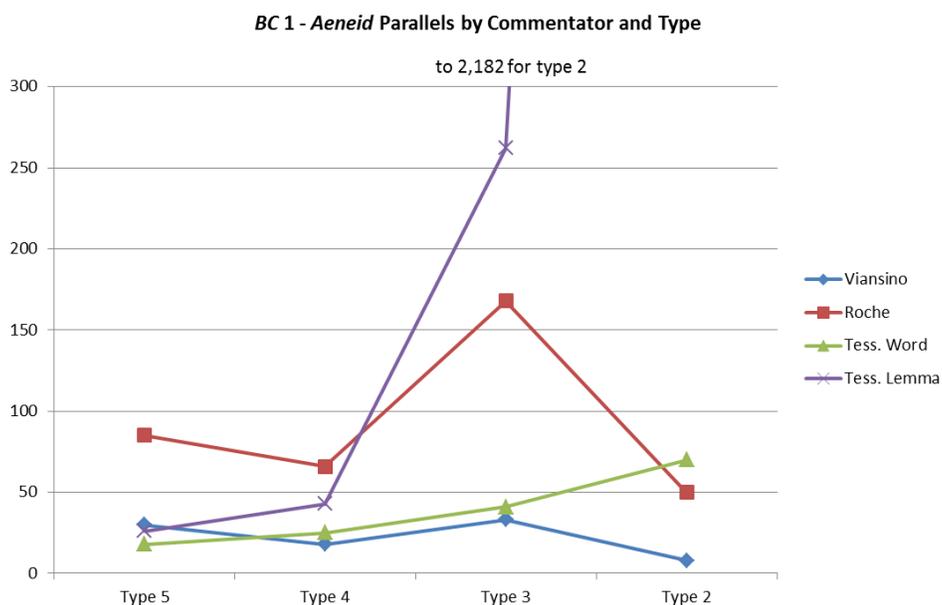


Note the expansion of the segments for types 3 and 2 as we go from Viansino to Roche, indicating the recovering of more common phrases and ordinary language, possibly as a result of electronic search.

The trend toward more parallels of types 3 and 2 is still more evident in the Tesseractae word and lemma search. Tesseractae lemma matching by whole sentence naturally returned far

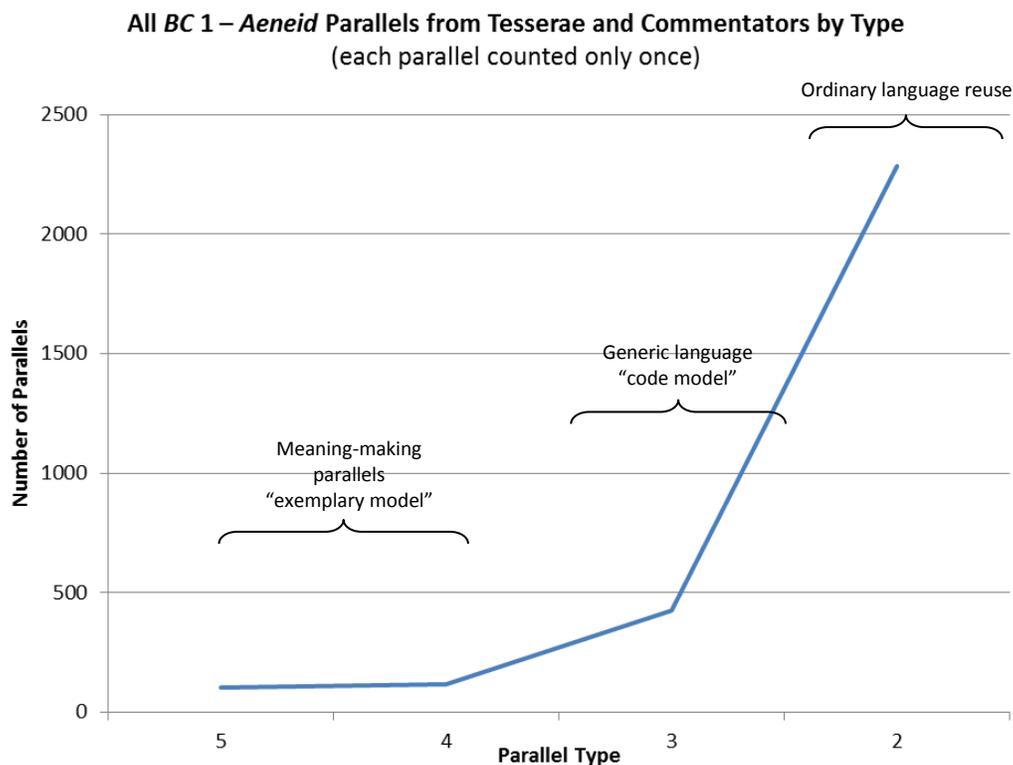
³⁸ Roche *per litteras* writes that the main electronic tool he used in researching the commentary “was the Packard Humanities Institute disk of Latin texts . . . using Musaios, typically for paired, rare or juxtaposed words, or for particular clausulae.”

more total candidates (2,994) than exact word search within six-word windows (159). Their distribution by type, along with those of Viansino and Roche, are shown in the following chart.



Beyond showing the increase in type 3 and 2 parallels with forms of electronic search, these results allow us to draw a further conclusion. The upward slanting lines in the two forms of Tesseræ search indicate that automatic detection of two-word similarity returns a high proportion of ordinary language. The commentators naturally filter most of this out, so that, unlike Tesseræ search, they return a much lower proportion of type 3 and type 2 parallels. Even if the commentators are willing to admit more common phrases and ordinary language when using electronic search, and so expand the definition of intertextuality in this way, better replication of their results will nonetheless require the ability to sort parallels automatically to allow for selective focus by type.

If rather than considering the differences between the sources, we combine them, we arrive at the best overall picture of *Aeneid* intertexts by type in *BC 1*.



Some of Lucan’s *Aeneid* intertexts originate with Vergil, others Vergil himself adapted from the common stock of Latin epic, while others are simply elements of the Latin language selected for metrical tractability. Distinguishing precisely between these three groups would require further investigation, but the distribution of our types 5-3 offers a useful approximation. Lucan has 218 interpretable (type 5 and 4) parallels to the *Aeneid* in *BC* 1, as compared to 425 instances of type 3 parallels. In the terms offered by Conte, this means that Lucan creates “code model” intertexts roughly twice as often as “exemplary model” intertexts. He uses ordinary (metrically compatible) language some 100 times as often. This distribution of intertext types generally accords with Zipf’s Law for natural language, which holds that the most common words in a language occur vastly more frequently than others, making for a roughly asymptotic distribution curve.³⁹ Although it makes intuitive sense that there will be far fewer meaningful intertexts than

³⁹ Zipf 1949. Recent discussion in Sampson 2003.

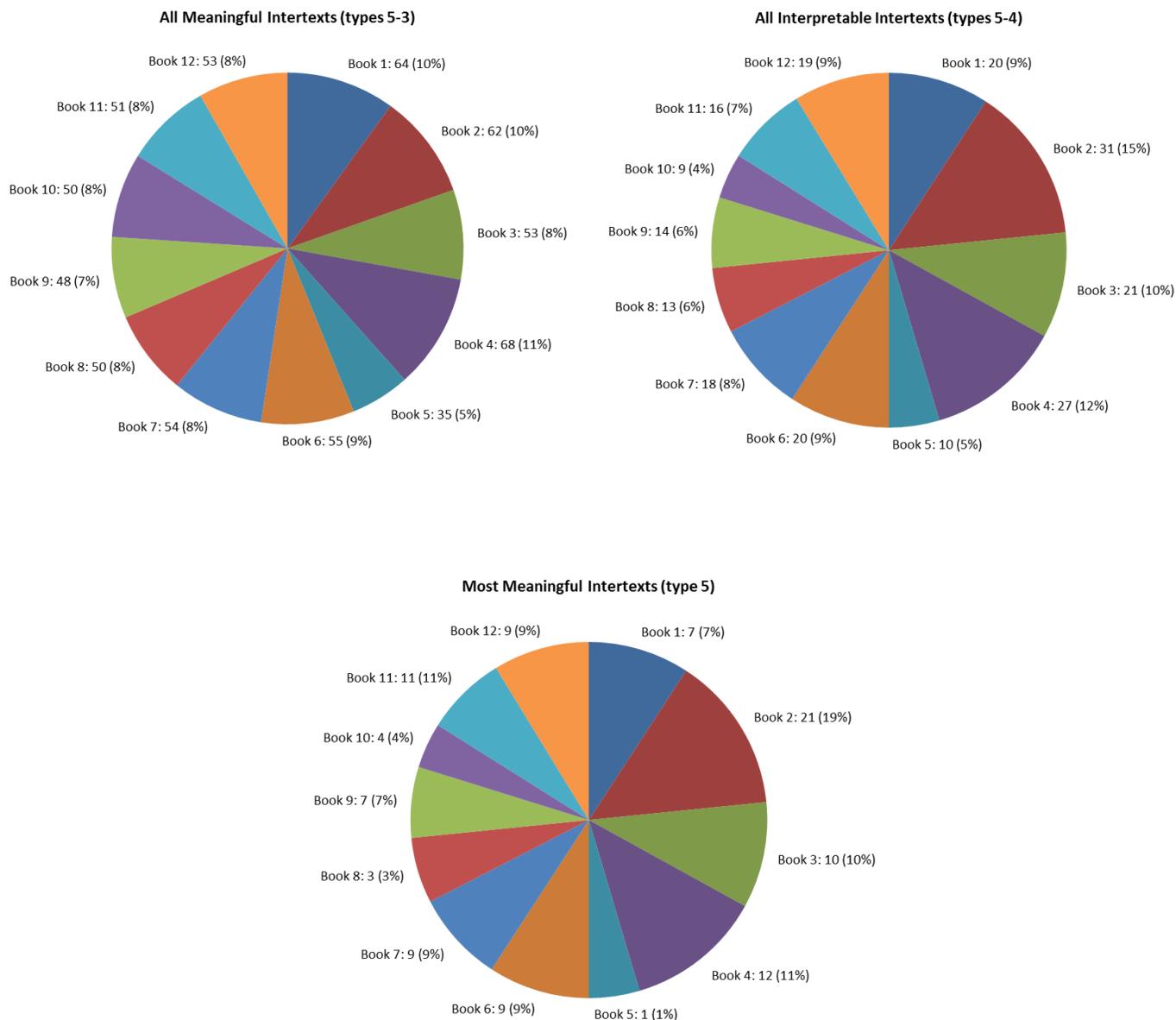
instances of ordinary (metrically compatible) language, it will require further study, including advances in automatic detection and classification, to determine whether this pattern in fact holds true for intertextual relationships generally. Such study can also illuminate the further question, more compelling for literary scholars, of how types of intertexts are distributed within an author's works and among different authors.

Analysis by Location in Source Text

The discussion of occurrence of intertexts by type has brought us to consideration of what combined information from commentators and automatic detection can say about the relationship between Lucan *BC* 1 and the *Aeneid*. Here we set out to answer three further questions: 1) which books of the *Aeneid* does Lucan draw from most in *BC* 1?, 2) How are his references distributed within *BC* 1?, and 3) how do these conclusions affect existing interpretations of the relationship between *BC* and the *Aeneid*?

The first of these questions was addressed above by looking at the *BC* 1 intertexts assigned by various sources to books of the *Aeneid*. Combining these sources and indicating distribution by type gives us the following results.

Lucan *Civil War* 1 Intertexts by Book of *Aeneid*



For each chart segment, we list in order the *Aeneid* book number, the number of intertexts in *BC* 1 from that book, and the percentage of total references this number constitutes. The first chart, representing Lucan's references of types 5-3, shows that in aggregate the poet draws fairly evenly on all books of the *Aeneid*, though he seems somewhat less interested in the games of

Aeneid 5, which gets only 5% of his references, and somewhat more in books 1 (10%), 2 (10%), and 4 (11%). Greater differences emerge when we consider only Lucan's use of interpretable references (types 5-4), illustrated in the second chart. For this purpose, Lucan clearly favors the first half of the *Aeneid*: references to the first half of the *Aeneid* make up 60% of the interpretable parallels in *BC* 1. If we proceed to consider only the most meaningful parallels, represented by the third chart, we find a still stronger preference for phrases from *Aeneid* book 2, as well as to a lesser degree for those from books 3, 4, 6, 7, 11, and 12.

To a certain extent, Lucan's preferences in *BC* 1 mirror those of ancient and modern audiences. We noted above the greater attention given by Servius to the first half of the *Aeneid*. Augustine was enthralled by the story of Dido. Books 2, 4, 6, and 12 of the *Aeneid* are standard parts of modern school curricula because the fall of Troy, the Dido story, the underworld, and Aeneas's final confrontation with Turnus remain among the most exciting and engaging parts of the poem. It could be that Lucan too found these the most compelling parts of the poem and thus had them closest to mind.

Closer scrutiny of context suggests more specific reasons for Lucan's choices. We might have conjectured that the Neronian poet would draw the greatest number of his parallels from *Aeneid* 2. In *BC* 1 he sets out his major theme of the Roman civil war as the destruction of Rome, for which he naturally draws upon Vergil's descriptions of the fall of Troy. Thus we find parallels such as *BC* 1.24-6 *semirutis . . . tectis / ingentia . . . / saxa iacent* \approx *Aeneid* 2.489 *pauidae tectis matres ingentibus errant*, where the tottering walls of Italian cities desolated by civil war recall the terror in Priam's palace at the invasion of the Greeks.

Lucan's reliance on *Aeneid* 4 seems a less obvious choice. Lucan adapts language from *Aeneid* 4 for several purposes, but a significant subset of his type 5-4 parallels (8 of 27) borrow

notions of fear and mania from the Dido story to show the same conditions gripping Rome before the outbreak of civil war. In one of these instances, Lucan refers to Vergil's combination of these two motifs: *BC* 1.676 *attonitam rapitur matrona per urbem* ≈ *Aen.* 4.666 *concussam bacchatur Fama per urbem*.⁴⁰ Lucan implicitly compares the frenzy of a matron predicting civil war to the dark rumors and madness that grip Carthage in the wake of Dido's death. Added to this are other allusions that suggest ill-starred similarities between Carthage and Rome, such as when Caesar casts off Roman law (*procul hinc iam foedera sunt*, *BC* 1.226) in a way that recalls Dido's denial of any treaties between Carthage and a future Rome (*nec foedera sunt*, *Aen.* 4.624). This complex of allusions to *Aeneid* 4 in *BC* 1 suggests that Romans of the civil war period were as mad, rumor driven, and (in the case of Caesar) lawless as passionate Dido and her Carthaginians ever were.

Lucan makes equally specific use of other *Aeneid* books. *Aeneid* 1 and 7 respectively initiate the poem and initiate its major war narrative. Lucan does both at once in *BC* 1, borrowing from *Aeneid* 1 and 7 to create resonances of poetic and martial beginnings.⁴¹ Lucan uses *Aeneid* 6 to show how the crime and corruption Vergil sequesters in the underworld is rampant in the real world of Roman affairs.⁴² *Aeneid* books 11 and 12 provide Lucan with military language, showing through premonitions that Rome's civil war will be as dire as the full blown battle of the Latins and Trojans.⁴³

⁴⁰ "The [prophetic] matron is swept through the city" ≈ "Rumor runs riot through the stunned city." The other instances are: *BC* 1.469 ≈ *Aen.* 4.173; 1.472 ≈ 4.183; 1.472 ≈ 4.189; 1.495-6 ≈ 4.300-1; 1.674-6 ≈ 4.300-1; 1.676 ≈ 4.68; 1.678 ≈ 4.110.

⁴¹ *Aen.* 1 parallels begin already with *BC* 1.2 *canimus* ≈ *Aen.* 1.1 *cano*. For *Aen.* 7, note e.g. *BC* 1.68 *immensum aperitur opus* ≈ *Aen.* 7.45 *maius opus moueo*; 1.311 *longa . . . pace* ≈ 7.47 *longa . . . pace*; 1.8 *licentia ferri . . .* 7.461 *amor . . . ferri*; 1.62 *limina Iani* ≈ 7.610 *limine Ianus*.

⁴² E.g. *BC* 1.13 *heu quantum terrae potuit* ≈ *Aen.* 6.828 *heu quantum inter se bellum*; 1.178 *rapti fasces pretio* ≈ 6.622 *fixit leges pretio atque refixit*; 1.681 *tela manusque* ≈ 6.57 *tela manusque*, with the reversal 1.256 *hac iter est bellis* ≈ 6.542 *hac iter Elysium nobis*.

⁴³ E.g. from book 11: *BC* 1.237 *clangorque tubarum* ≈ *Aen.* 11.192 *clangorque tubarum*; 1.388 *it . . . clamor* ≈ 11.192 *it clamor*; 1.4-5 *rupto foedere regni / certatum totis concussi uiribus orbis* ≈ 11.313 *toto certatum est*

The discovery of a number of significant parallels to the relatively understudied book 3 by all sources suggests the sensitivity of each method to Lucan's artistry apart from his reception. A significant strand in Lucan's use of *Aeneid* 3 involves reversing that book's optimistic prophecies of a new Trojan homeland to suggest the woeful future in store for the Roman people.⁴⁴ One parallel discovered only by Tesseract search exemplifies this motif. At *Aeneid* 3.418, the seer Helenus advises Aeneas to avoid Scylla and Charybdis, and notes that their domain along the sea is where Sicily separated from Italy, *Hesperium Siculo latus abscondit*. Vergil's image is one of opening up the straits of Messina for Aeneas's voyage. At *BC* 1.547, Lucan uses the language of Vergil to close the same sea passage. As a portent of the coming civil war, Mt. Aetna erupted and rejoined Sicily and Italy: *Hesperium cecidit latus*. Whereas the phrase of Helenus suggested the unlocking of possibilities in the new Italian world, Lucan's reversal describes a calamity signalling the end of the Roman power Aeneas had established.⁴⁵

The attention given to these books contrasts with Lucan's relative neglect of others. Lucan draws least across all types from *Aeneid* book 5, and little from book 8. We may surmise that the relatively tranquil subject matter of these books, including the Trojan games and Aeneas's sojourn with Evander, bore too little relation to Lucan's story of Roman catastrophe. Exceptions show the rule, as when Lucan borrows Vergil's phrase for the cold blood of the boxer Entellus to describe the hot blood of Caesar's troops (*Aen.* 5.395 *gelidus . . . / sanguis* \approx *BC* 1.363 *calidus . . . sanguis*). Where Lucan can thus intensify language from the *Aeneid* in the

corpore regni. From book 12: 1.157 *dat stragem* \approx 12.454 *dabit . . . / stragemque*; 1.366 *usque adeo miserum est ciuili uincere bello* \approx 12.646 *usque adeone mori miserum est*; 1.388 *it . . . clamor* \approx 12.409 *it clamor*.

⁴⁴ E.g. *BC* 1.196 *Phrygiique penates* \approx *Aen.* 3.148 *Phrygiique penates*, 1.224 *Hesperiae . . . aruis* \approx 3.171 *Dictea negat tibi Iuppiter arua*; 1.524 *manifesta fides* (of ills to come) \approx 3.375 *manifesta fides* (of divine guidance). O'Hara 1990 shows that such prophecies are in fact already proved false in the *Aeneid* itself.

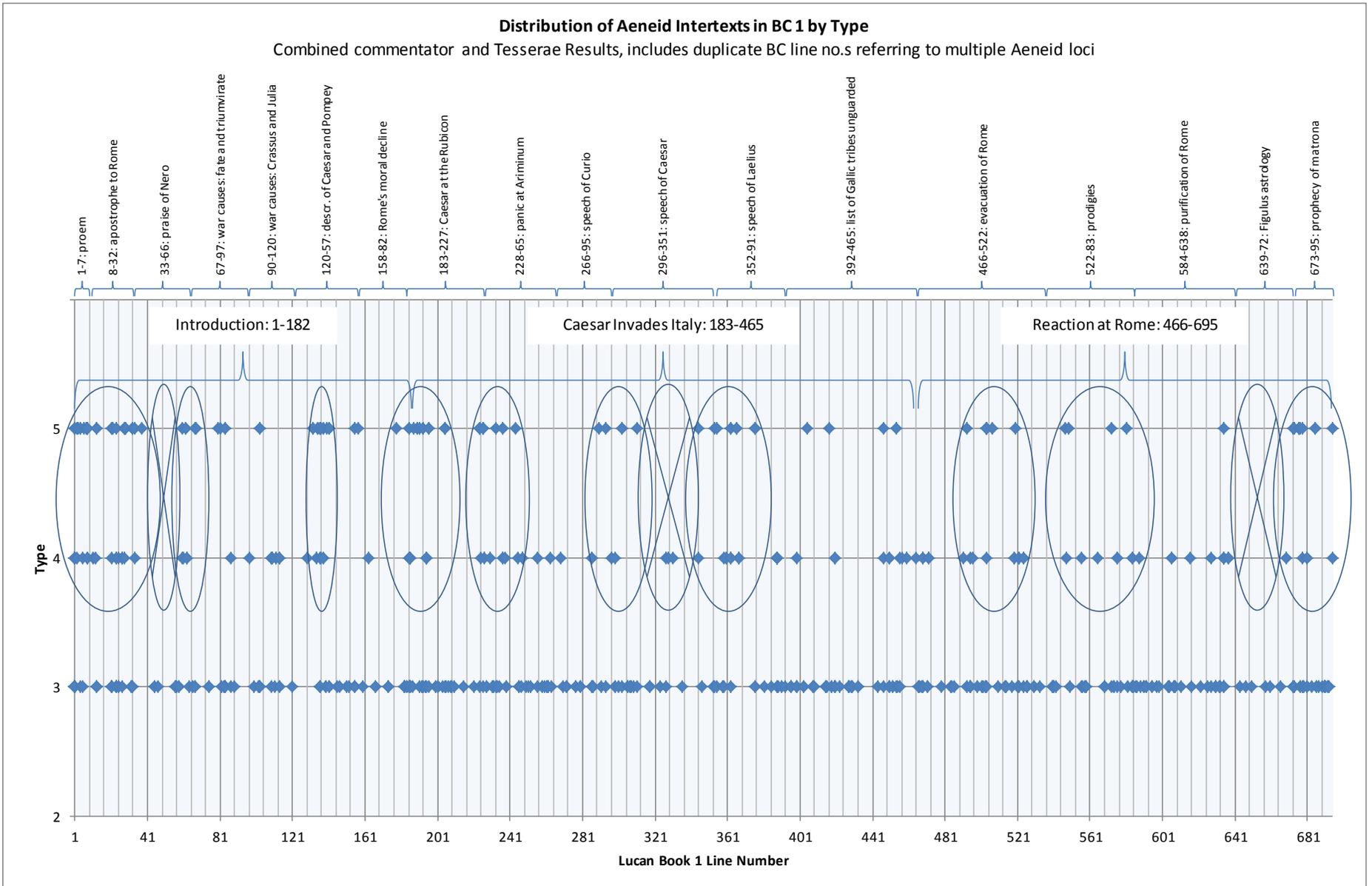
⁴⁵ Lucan is reversing Ovid as well, though without the verbal parallels. Ovid *Met.* 15.290-2 refers to accounts that Sicily was separated from the mainland by a surge of the sea.

direction of conflict, he seizes the opportunity, but otherwise passes over Vergil's tenuous discourses of normalcy and harmony.

The poetic effects of Lucan's intertexts with individual books of the *Aeneid* can all be (and have been) investigated further. We hope this survey has sufficed to shed some light on this aspect of Lucan's artistry and to demonstrate that large-scale analysis of intertextual connections by source location can be productive.

Analysis by Location in Target Text

We offer a last perspective on large-scale intertextuality by considering the distribution of *Aeneid* intertexts within *BC* 1.



This chart shows the locations of parallels to the *Aeneid* in *BC* 1 by type. The x-axis shows line numbers in *BC* 1, and the y-axis the type of parallel (only those of types 5-3 are indicated). Each dot represents one parallel. When one locus in *BC* echoes more than one in the *Aeneid*, more than one dot appears at the relevant *BC* locus in the chart. The steady pattern of type 3 parallels shows Lucan drawing on the *Aeneid* for epic (“code model”) diction and metrically compatible phrases evenly throughout *BC* 1. The consistency of Lucan’s use of Vergil at this level suggests the near inevitability for the Neronian poet of using the epic materials Vergil provided.

The distribution of interpretable parallels in *BC* 1 will be of greater interest. We have placed ovals around the most significant groupings, and crossed-out ovals in areas completely or relatively lacking in parallels. Among types 4 and 5, the most conspicuous pattern is the thick cluster at the opening of the poem, in the proem at lines 1-7 and in the apostrophe to Rome at lines 8-32. Scholars have demonstrated how in these lines allusions to the *Aeneid* and *Iliad* identify Lucan’s poem as a new sort of epic focusing on repetitive destruction rather than linear progress.⁴⁶ Another cluster in the opening descriptions of Caesar and Pompey in lines 120-57 links each with figures in the *Aeneid* and epic tradition. A final group, in the prophecy of the matrona, closes the book. Lucan thus shows a pattern of employing interpretable references in opening and closing sections, as well as in a highly significant passage where he establishes his principal characters.

If we look more closely at the group of intertexts that open the book, we find a notable absence. Within a thicket of Vergilian language, references to the *Aeneid* suddenly cease in the middle of Lucan’s praise of Nero, lines 39-59. Roche has noted that:

⁴⁶ Overview at Roche 2009: 22-23.

the invocation [to Nero] is in a radically different mode from the epic announced at 1-32. It raises questions of propriety of genre, made more urgent by the cataclysmic subject matter which Nero is said by Lucan to inspire.⁴⁷

The change can be explained in part by Lucan's shift to the *Georgics* as a model for his panegyric.⁴⁸ But the cessation of *Aeneid* intertexts at the height of his praise of Nero also means that Lucan foregoes an opportunity to associate Nero with the great mythical heroes of epic in favor of references to real-world Rome, creating a more prosaic tone that tempers his encomium. In the other major lacuna in *BC* 1 at lines 639-72, Lucan ceases significant references to the *Aeneid* during the astrological forecast of civil war by Nigidius Figulus.⁴⁹ Again the poet turns to the *Georgics* as well as Lucretius and other didactic writers.⁵⁰ The silencing of high epic resonances seems consonant with scientific prophecy in a way that is not with praise of a *princeps*.

In addition to sections of concentrated references and lacunae, we also find a tendency to cluster references at the beginning and ending of sections. As Lucan moves from the praises of Nero to the causes of war (61-68), from the causes of Rome's decline to Caesar's halt at the Rubicon (178-205), from the Rubicon to panic at Ariminum (224-44), from Curio's speech to Caesar's (290-311), and from Caesar's speech to that of Laelius (345-76), he bridges his sections with some of his most notable and significant references to the *Aeneid* (each of these sets are circled in the chart above). By clustering Vergilian references around section openings and closings, Lucan brings brings density of meaning and Vergilian authority to his transitions.

⁴⁷ Roche 2009: 9.

⁴⁸ Particularly *G.* 1.24-62. Roche 2009: 9.

⁴⁹ Casali 2011: 92-95, which was published after our testing, suggests several parallels between this episode and the *Aeneid*. Only one of these, cited on p. 95, meets the criteria for an interpretable parallel in our classification scheme. This is *BC* 1.649-50 *quod cladis genus, o superi, qua peste paratis / saevitiam?* ≈ *Aen.* 2.361-2 *quis cladem illius noctis, quis funera fando explicet . . . ?* We would rank the parallel as a type 5, so in fact we have one intertext recalling the destruction of Troy as a precedent for the calamity at Rome to give political point to the portents of nature Figulus describes.

⁵⁰ E.g., Roche 2009: 370-371 cites *BC* 1.657 ≈ *Lucr.* 2.413, *Man.* 5.744 f., as well as *Sen. Her. O.* 1385-7; 1.658-9 ≈ *Verg. G.* 1.34f.

Within these transitions, Lucan employs a specific pattern that highlights his own virtuosity. Just prior the close of a section, he presents his audience with a Vergilian intertext, then finishes with his own novel *sententia*, before beginning a new section freshly anchored and authorized by references to the *Aeneid*. Thus toward the end of the enumeration of the causes of Rome's decline, we find the bribery of the Vergilian underworld translated into the real world of Lucan's Rome (*BC* 1.178 *rapti fasces pretio* \approx *Aen.* 6.622 *fixit leges pretio atque refixit*). Lucan then closes his accounting with a sententious vision of avidity leading to war (*concussa fides et multis utile bellum*, 1.182). He then opens the section on Caesar at the Rubicon with several references to the *Aeneid*. In one, noted above, the specter of the city of Rome (1.185 *ingens . . . patriae trepedantis imago*), revealing Caesar's betrayal of *pietas*, recalls Aeneas's thoughts of his relationship with his father after unwillingly killing the boy warrior Lausus (10.824). In another, Roma calls out to Caesar's troops in a way that suggests they are as confused as disheartened Trojans fleeing battle rather than coming to it justly (*BC* 1.190 *quo tenditis ultra?* \approx *Aen.* 9.781 *quo deinde fugam, quo tenditis?*). Lucan's transition from the speech of Caesar to that of Laelius follows a similar pattern.⁵¹ Here and elsewhere, Lucan uses Vergilian authority and color to hold and excite his audience during scene transitions, but reserves section endings for his own master strokes.

The grouping of intertexts around transitions can also create the sort of register changes seen in the praise of Nero and the prophecy of Nigidius Figulus. We find *Aeneid* intertexts in the transitions to and from Caesar's first speech to his troops. Caesar begins in the grand tones of Aeneas consoling his men (*BC* 1.299 *bellorum o socii, qui mille pericula Martis . . .* \approx *Aen.* 1.198-9 *o socii (neque enim ignari sumus ante malorum)*), while denigrating the life of peace

⁵¹ *BC* 1.345 \approx *Aen.* 3.85; sententious ending (*detrhimus dominos urbi seruire paratae*, 1.351); further significant *Aeneid* references at *BC* 1.354, 355, 363, 366.

which in the *Aeneid* seems a lost ideal (1.311 *ueniat longa dux pace solutus* \approx 7.46-7 *urbes* / . . . *longa placidas in pace regebat*). He ends with the tendentious reference to city walls owed to his men considered above (1.345 \approx 5.717). Vergilian intertexts are not entirely absent from the middle of the speech: Caesar borrows Vergilian language to compare Pompey to a bloodthirsty tiger (1.328 *matrum* . . . *lustra* \approx 3.646-7 *ferarum* / *lustra*). But this parallel is of a different order from his invocation of grander Vergilian subjects. At the heart of his appeal, Caesar descends from epic grandeur into petty deprecation of Pompey (314-40) and complains about his lost gains (340-2). Caesar's intertextual slander of Pompey as no more than a savage animal contributes to this shift in register and the impression that Caesar's nobler opening and closing sentiments are just a veneer.

BC 1 and the Aeneid

Having explored two large-scale views of the intertextual relationship between *BC 1* and the *Aeneid*, we can now ask what these perspectives add to our understanding of Lucan's epic. In terms of the artistic relationship between the two poets, Lucan's consistent use of Vergilian language pays tacit tribute to his predecessor, and his careful deployment of *Aeneid* intertexts in section transitions attests to the power he saw in the use of Vergilian precedent. In thematic terms, Lucan's references to the *Aeneid* have generally been taken as oppositional: he subverts passages from the *Aeneid* to suggest that the construction of empire, whose costs and benefits Vergil carefully balances, was inevitably corrupt. An uncontrollable thirst for power destroyed pristine Roman values of loyalty to family, community, and country.⁵²

⁵² Syndikus 1958: 89-90 characterizes *BC* as a "Gegenbild" of the *Aeneid*. Roche 2009, gives an overview of this line of thought, writing of the "ideological opposition to the surface narrative of the *Aeneid*" found in *BC*. The most recent treatment is Casali 2011.

A large-scale view allows us to see how this oppositional tenor was articulated in *BC 1* through a series of local strategies. Beyond amplifying and trumping the opening of the *Aeneid* in the first book of his own epic, Lucan creates a series of transformative identifications. Civil war recapitulates the destruction of Troy (*Aen.* 2). Prophecies of progress (*Aen.* 3) turn to forebodings of catastrophe. The madness and panic of Carthage (*Aen.* 4) spreads to all of Rome. Crimes worthy of the underworld (*Aen.* 6) appear on earth. Fate as a guiding force in war (*Aen.* 7) is shown hollow.⁵³ The purpose of military honor, even as hesitantly posed by Vergil (*Aen.* 11), is questionable.⁵⁴ Caesar is assimilated to Turnus while Pompey appears weaker than Aeneas (*Aen.* 12).⁵⁵ These are just some of the ways in which Lucan refers to individual books of the *Aeneid* to demonstrate that civil war is the latest and most devastating repetition of previous mythical struggles, and that Vergil's honorable mythology of Rome's foundation is but a shadow of the truth.

Digital means are not of course the only path to such large-scale understandings of intertextuality. We have given but a miniature version of the vast comparisons carried out manually by Knauer and Nelis. The point is rather that digital detection can accelerate large-scale analysis considerably, and digital manipulation of results allows for multiple ways to view intertextual relationships and develop meaningful generalizations about them.

Theoretical Consequences and the Future of Intertextuality

⁵³ *BC* 1.34 *inuenere uiam* ≈ *Aen.* 7.297 *inuenere uiam*; 1.635 *di uisa secudent* ≈ 7.259 *di nostra . . . secudent*.

⁵⁴ The cowardice of Drances becomes the greater concern of fields left vacant by war (*BC* 1.29 *desuntque manus poscentibus aruis* ≈ *Aen.* 11.379 *bella manus poscunt*); the narrator asks why Rome rushed to war, in contrast with the Latin leader Tarchon who asks why his troops will not fight (1.8 *quis furor o ciues quae tanta licentia ferri?* ≈ 11.732 *quis metus, o numquam dolituri, o semper inertes / Tyrrheni, quae tanta animis ignavia venit?*).

⁵⁵ Caesar-Turnus: *BC* 1.205-12 *leo...///...erexitque iubam.../...infremuit...lancea* ≈ 12.4-9 *leo...comantis / excutiens ceruice toros.../...telum et fremit*. Pompey weaker than Aeneas: 1.134 *nouas uires* ≈ 12.424 *nouae . . . uires*.

Among the most significant consequences of the development of digital search is the potential for more precise definitions of intertextuality. Consider the following table, which lists numbers of type 5-4 parallels identified in our combined tests (both Tesseract searches and all commentators) grouped by features.

Minimal Features of Type 5-4 BC 1 – *Aeneid* Parallels

| | Features | Count | % of total |
|----|---|-------|------------|
| 1. | Two-word (exact word or lemma) identity | 146 | 67% |
| 2. | One identical word + semantic context | 27 | 12% |
| 3. | One identical word + synonym | 16 | 7% |
| 4. | Semantic context only | 13 | 6% |
| 5. | Synonyms only | 12 | 6% |
| 6. | One identical word + syntax | 2 | 1% |
| 7. | One identical word + sound | 2 | 1% |
| | TOTAL | 218 | 100% |

Tesseract looked only for parallels containing at least two similar words, those in the first category above. But every parallel identified by Tesseract or commentators could have been found by a search sensitive to the seven feature categories listed. Two-word similarity was sufficient to identify 67% of the parallels identified by all sources. A search for single identical word in a highly related context would have found 12%, a search for a single identical word and synonym, 7%. A search for similar thematic material, without words in common, would have found 6%, a search for paired synonyms another 6%, and so forth. Of these features, one is already subject to automatic detection with the Tesseract tool (word identity). Linguistic techniques for identifying semantically related words and word clusters could be employed to identify synonyms and antonyms and to identify thematic contexts.⁵⁶ Methods for syntax

⁵⁶ Latent semantic analysis, a technique developed by linguists to find words with associated meanings, is increasingly being used by humanists to search for semantically related words and repeated themes. See Underwood 2011. Topic modeling is a related semantic detection technique that identifies themes in the form of clusters of co-

matching have been demonstrated in intertextual analysis,⁵⁷ and sound matching techniques are in development.⁵⁸

Full sensitivity to these seven features and feature combinations thus holds the potential to capture all of the interpretable intertexts found by commentators in our sample and others in addition. To build this search configuration into a full algorithm for intertextual analysis would require the addition of at least two other capacities, however. One is the ability to sort parallels into types. The table above accounts only for our type 5-4 parallels, which we selected manually from the total Tesseract and commentator results. We are currently exploring whether word frequency, phrase frequency, and other feature matching may be helpful in creating an automatic ranking system. Another necessary expansion will be the development of cross-linguistic matching. For the study of Latin intertextuality, a Greek-Latin matching algorithm should at a minimum recognize parallelism such as Homer *Od.* 1.1 ἄνδρα (“man”) ≈ *Aen.* 1.1 *uirum* (“man”). Such similarity would seem to be detectable: cross-linguistic semantic and morphological (for case) matching along with sensitivity to section boundaries would rate these words highly as a potential parallel.⁵⁹ As such methods are developed, they can be tested on an expanding range of primary texts and their results compared with those from the relevant secondary sources. This process seems likely to result in algorithms sensitive to the full range of intertextuality as recognized by critics.

Once an algorithm has reached a satisfactory level of performance in sample tests in terms of recall (capturing known parallels) and specificity (sorting high and low interest

occurring words. It has been used, for example, to identify themes such as midwifery, church, death, gardening, shopping, and illness in the diary of Martha Ballard, a midwife living in late 18th – early 19th century Maine (Blevins 2010).

⁵⁷ Bamman and Crane 2008.

⁵⁸ Forstall and Scheirer 2010.

⁵⁹ Work on cross-linguistic semantic matching has included alignment of translations with target texts to find translation equivalents. See Bamman, Babeu et al. 2010, available at <http://www.perseus.tufts.edu/hopper/about/publications>.

parallels), it can be set to work on comparisons of many individual texts with some confidence in the results it will return. The full set of comparisons would then provide a baseline against which to measure the intertextual activity of genres, authors, works, and work sections synchronically or over time. We might thus think of the table above as a first version of an intertextual “fingerprint” that in the future can be compared with fingerprints for other authors and texts to arrive at a better idea of how each author and text are distinctive and of overall intertextual practice.

This program of intertextual study presents us with a theoretical question: if an algorithm identifies substantially all intertexts found by critics, does it constitute a definition of intertextuality? We would answer yes, provided that we distinguish between a definition and a description. The sort of detection algorithm we have outlined cannot describe the whole phenomenon of intertextuality any more than a recipe for apple pie describes the experience of baking and eating one. As discussions of Latin intertextuality from Pasquali onward have shown, intertextual phenomena show a range of subtlety in the connections they establish and meanings they form. Even if search algorithms proceed further into the detection of Alexandrian footnotes, window references, structural similarities, and other complex phenomena, the definitional truth they represent will always need to be complemented by the experiential truth of the reader to form a fully satisfying description.

Nevertheless, a definition can be a useful complement to such description. A definition identifies only the minimal features necessary to distinguish one thing from another. The most minimal definition of intertextuality might be “similarity between any texts.” To make such a definition more useful, however, we must make it more specific. Here we can turn for guidance to the intuitions of scholars well versed in intertextual artistry. In a discussion of the problem of

authorial intentionality, Joseph Farrell appeals to the statistical unlikelihood of the random production of two very similar lines,

Ἄλκανδρόν θ' Ἄλιόν τε Νοήμονά τε Πρύτανίν τε (*Il.* 5.678), and

Alcandrumque Haliumque Noëmonaque Prytanimque (*Aen.* 9.767).

and makes this unlikelihood the ultimate justification for declaring the lines related.

One does have to be a qualified reader to grasp the fact that Vergil's line is intertextual with Homer. But once this relationship is grasped, one may ask whether the similarity exists only in the mind of the reader and not at all in the text or in the mind of the poet. I am not a statistician, but I think we can assume a very low mathematical probability that any line in any Latin poem would, just by chance, reproduce so accurately any line from any Greek poem. That is to say, the exactitude of this similarity, and the unlikelihood that such a similarity would occur by chance—neither of which factors is merely impressionistic—are both so great as to make the line itself an unmistakable intertextual marker.⁶⁰

Farrell is arguing for the improbability of Vergil unintentionally producing such a close echo of Homer. To substantiate this argument, he needs to make the case that Vergil's line obviously *is* intertextual with the line of Homer, and would have seemed as obviously so to Vergil as it does to us. When he looks for the strongest argument that this is in fact an intertext, Farrell turns to formal and quantitative features. That is, when looking for the soundest definition of intertextuality he can find, Farrell turns to an algorithm. In his case, it is a hypothetical algorithm that accounts for the probability of a semantic and sound convergence between lines of Homer and Vergil.

If we pursue Farrell's reasoning further, we can make this definition yet more precise. His argument suggests a notional spectrum, where parallels with rare configurations of similarity are placed at one end and common co-occurrences at the other. We could then place a line on the spectrum separating intertexts from non-intertexts. In its fullest and most satisfying form, the "intertext" side of the spectrum would consist of a composite of all combinations of features—

⁶⁰ Farrell 2005: 100-101.

the semantics and sound that Farrell addresses, along with others—that qualified a proposed parallel as an intertext. Better still, rather dividing the spectrum into two parts, we could segment it into types, where Farrell’s Homer-Vergil parallel would appear among interpretable intertexts at one end. The full, segmented spectrum, or an algorithm that produced it, could then be understood as an empirical definition of intertextuality, much as animal species are defined by a set of morphological and behavioral features.

As described, such a definition would largely replicate the criteria critics have traditionally used to describe intertexts. The process of digital discovery can also expand our definition, however. To a degree, this is already happening. Our comparison of commentators showed that those who use digital detection methods find more instances of pure language reuse (type 3). The collective scholarly understanding of intertextuality is thus shifting to include phenomena considerably less salient than the most celebrated examples. Digital detection may also bring to prominence other features that have passed beneath the notice of critics, but which have nevertheless played a role in the poet’s production and reader’s experience. When we listen to music, we respond not only to melody and rhythm, but also to more obscure elements such as structure, overtones, and established expectations. We have argued for a similar effect when Lucan silences the otherwise steady drone of foreground (type 5-4) and background (type 3) intertexts with the *Aeneid*. Such arguments could be made for sound patterns and other features.⁶¹

To be sure, digital detection will contribute to forming not one definition of intertextuality, but many. Substituting other features for the seven we identified might create another definition that was equally comprehensive. Our results were dependent upon a particular

⁶¹ T.S. Eliot writes “I cannot help suspecting that to the cultivated audience of the age of Virgil part of the pleasure of poetry arose from the presence in it of two metrical schemes in a kind of counterpoint, even though the audience may not have been able to analyse the experience” (Eliot 1942: 12, cited by Wilkinson 1970: 120).

set of commentators and sample texts. Other combinations might produce definitions with other features. It is not unreasonable to expect, however, that from such investigations an algorithm accounting for substantially all levels of intertextuality could emerge.⁶²

By offering a definition of intertextuality, we have perforce entered into the theoretical discussions of recent years over the status of intertexts.⁶³ In these discussions, three main alternatives have been on offer: intertexts exist are created by an author; they are features of texts; or they are constituted by an observer connecting two texts while reading or listening. Digital detection and analysis do not change the terms of this debate. But they can bring an increased level of precision to the positions.

A search algorithm might be construed as simply an extension of the reader, in the form of a systematic projection of a set of reader preferences and expectations. From this vantage point, our efforts can be seen as an attempt to reverse-engineer an intertextual detection algorithm latent in the mind of the critic. However we might conceive of this operation, making readerly criteria clear for all to see allows them to be compared, contested, and refined. Alternatively, digital search may be the ultimate means of focusing on the text alone. The fact that our algorithm identifies unobserved instances and minimally-recognized features of intertextuality suggests that it removes at least one layer of subjectivity by presenting to the critic a set of examples that all fit consistent criteria. Finally, although we cannot recover the author's intention, developing a stylistic fingerprint for each author can fill in the picture of an artist's preferences and working methods. Critics have long observed that Ovid employs dactyls in the *Metamorphoses* more often than Vergil does in the *Aeneid*. Reading this difference together with

⁶² Lausberg 1985: 1616-1617 argues that Homer and Vergil had such weight in the epic tradition that no lexical similarity is necessary to invoke comparison, an observation approved by Henderson 2010 [1987]: 479. Whether the precise levels of similarity Lausberg refers to are automatically detectable, or would be defined by most critics as intertextuality, requires further exploration.

⁶³ Farrell 2005 provides a useful entry point.

the content of the poems, we conclude that Ovid's swifter meter contributes to a lighter narrative style and tone.⁶⁴ If we are willing to accept a certain amount of distant reading, digital analysis can provide an analogous and complementary account of each poet's intertextual practice. When intertextuality, and potentially other complex poetic features, can be thus modeled, we can build a more satisfying profile of the artistry of each poet and richer interpretations of their works.

⁶⁴ 54.6% of Ovid's first four hexametrical feet are dactyls, to 44.1% in the *Aeneid* (Wilkinson 1970: 131, with further references). Wilkinson writes that Ovid's "fantastical stories could only bear the lightest touch."

Works Cited

- Bamman, D., A. Babeu, et al. 2010. "Transferring Structural Markup Across Translations Using Multilingual Alignment and Projection." *Proceedings of the 10th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2010)* 2010-6: 11-20.
- Bamman, D. and G. Crane. 2008. "The Logic and Discovery of Textual Allusion." *Proceedings of the Second Workshop on Language Technology for Cultural Heritage Data (LaTeCH 2008) (Marrakesh)*.
- Barchiesi, A. 1984. *La traccia del modello: effetti omerici nella narrazione virgiliana*. Pisa, Giardini.
- Blevins, C. 2010. "Topic Modeling Martha Ballard's Diary." *historying*
<http://historying.org/2010/04/01/topic-modeling-martha-ballards-diary/> Accessed February 26, 2012.
- Büchler, M., A. Geßner, et al. 2010. "Unsupervised Detection and Visualization of Textual Reuse on Ancient Greek Texts." *Proceedings of the Chicago Colloquium on Digital Humanities and Computer Science* 1.
- Casali, S. 2011. "The Bellum Ciuile as Anti-Aeneid". In *Brill's Companion to Lucan*. Ed. P. Asso. Leiden, Brill. 81-109.
- Conte, G. B. 1986. *The Rhetoric of Imitation: Genre and Poetic Memory in Virgil and Other Latin Poets*. Ithaca, Cornell University Press.
- D'Ippolito, G. 2000. "Il concetto di intertestualita nel pensiero degli antichi." *Classical Salmanticensia*: 13-32.

- De La Cerda. 1617. *P. Virgilii Maronis libri Aeneidos: argymentis, explicationibvs notis illvstrati*. Lyon, Sumptibus Horatii Cardon.
- Edmunds, L. 2001. *Intertextuality and the Reading of Roman Poetry*. Baltimore, Johns Hopkins University Press.
- Eliot, T. S. 1942. *The Music of Poetry*. Glasgow, Jackson, Son & Company.
- Fantuzzi, M. and R. Hunter. 2004. *Tradition and Innovation in Hellenistic Poetry*. Cambridge, Cambridge University Press.
- Farrell, J. 1991. *Vergil's Georgics and the Traditions of Ancient Epic: The Art of Allusion in Literary History*. New York, Oxford University Press.
- Farrell, J. 2005. "Intention and Intertext." *Phoenix* 59: 98-111.
- Forstall, C. and W. Scheirer. 2010. "Features from Frequency: Authorship and Stylistic Analysis Using Repetitive Sound." *Journal of the Chicago Colloquium on Digital Humanities and Computer Science* 1.
- Giangrande, G. 1967. "'Arte Allusiva' and Alexandrian Poetry." *CQ* 17: 85-97.
- Heitland, W. E. and C. E. Haskins. 1887. *M. Annaei Lucani Pharsalia*. London, G. Bell.
- Henderson, J. 2010. "Lucan/The Word at War". In *Lucan*. Eds C. Tesoriero, F. Muecke and T. Neal. Oxford, Oxford University Press. 433-492.
- Hinds, S. 1998. *Allusion and Intertext: The Dynamics of Appropriation in Roman Poetry*. New York, Cambridge University Press.
- Hinds, S. 2007. "Martial's Ovid/Ovid's Martial." *JRS* 97: 113-154.
- Horsfall, N. 2000. *Vergil, Aeneid 7: A Commentary*. Leiden, Brill.

- Horton, R., M. Olsen, et al. 2010. "Something Borrowed: Sequence Alignment and the Identification of Similar Passages in Large Text Collections." *Digital Studies / Le champ numérique* 2.
- Knauer, G. N. 1964. *Die Aeneis und Homer: Studien zur poetischen Technik Vergils, mit Listen der Homerzitate in der Aeneis*. Göttingen, Vandenhoeck & Ruprecht.
- Lausberg, M. 1985. "Lucan and Homer". In *ANRW*. Eds W. Haase and H. Temporini. Berlin-New York. 1565-1622.
- Levene, D. S. 2010. *Livy on the Hannibalic War*. Oxford, Oxford University Press.
- Martindale, C. 1993. *Redeeming the Text: Latin Poetry and the Hermeneutics of Reception*. Cambridge, Cambridge University Press.
- Moretti, F. 2005. *Graphs, Maps, Trees: Abstract Models for a Literary History*. London and New York, Verso.
- Nelis, D. 2001. *Vergil's Aeneid and the Argonautica of Apollonius Rhodius*. Leeds, Francis Cairns.
- O'Hara, J. J. 1990. *Death and the Optimistic Prophecy in Vergil's Aeneid*. Princeton, Princeton University Press.
- Pasquali, G. 1968. "Arte allusiva". In *Pagine Straviganti*. Ed. G. Pasquali. Florence, Biblioteca Sansoni. 275-283.
- Roche, P., Ed. 2009. *Lucan: De bello civili. Book 1*. Oxford, Oxford University Press.
- Sampson, G. 2003. "Statistical Linguistics". In *International Encyclopedia of Linguistics*. Ed. W. Frawley. Oxford, Oxford University Press. 162-165.
- Saylor, C. 1978. "Belli Spes Improba: The Theme of Walls in Lucan's Pharsalia VI." *TAPA* 108: 243-257.

- Schepens, G. and J. Bollansée. 2005. *The Shadow of Polybius: Intertextuality as a Research Tool in Greek Historiography: Proceedings of the International Colloquium, Leuven, 21-22 September 2001*. Leuven, Peeters.
- Swift, L. 2010. *The Hidden Chorus: Echoes of Genre in Tragic Lyric*. Oxford, Oxford University Press.
- Syndikus, H. P. 1958. *Lucans Gedicht vom Bürgerkrieg: Untersuchungen zur epischen Technik und zu den Grundlagen des Werkes*, Munich.
- Telo, M. 2010. "Embodying the Tragic Father(s): Autobiography and Intertextuality in Aristophanes." *Classical Antiquity* 29: 278-326.
- Thomas, R. F. 1986. "Virgil's Georgics and the Art of Reference." *HSCP* 90: 171-198.
- Thomas, R. F. 1999. *Reading Virgil and His Texts: Studies in Intertextuality*. Ann Arbor, University of Michigan Press.
- Thompson, L. and R. T. Bruère. 1968. "Lucan's Use of Vergilian Reminiscence." *CP* 63: 1-21.
- Underwood, T. 2011. "LSA is a marvellous tool, but literary historians may want to customize it for their own discipline." *The Stone and the Shell*
<http://tedunderwood.wordpress.com/2011/10/16/lisa-is-a-marvellous-tool-but-humanists-may-no-use-it-the-way-computer-scientists-do/> Accessed February 26, 2012.
- Viansino, G. 1995. *Marco Annaeo Lucano: La Guerra Civile (Farsaglia) libri I-V*. Milan, Arnoldo Mondadori.
- Wilkinson, L. P. 1970. *Golden Latin Artistry*. Cambridge, Cambridge University Press.
- Wills, J. 1996. *Repetition in Latin Poetry: Figures of Allusion*. Oxford, Oxford University Press.
- Zipf, G. K. 1949. *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. Cambridge, Mass., Addison-Wesley Press.