Measuring and Mapping Intergeneric Allusion in Latin Poetry using Tesserae

Patrick J. Burns

Journal of Data Mining and Digital Humanities

Measuring and Mapping Intergeneric Allusion in Latin Poetry using Tesserae

Patrick J. Burns[1] *

1 New York University, United States

*Patrick J. Burns: patrick.j.burns@nyu.edu

Abstract
Most intertextuality in classical poetry is unmarked, that is, it lacks objective signposts to make readers aware of the presence of references to existing texts. Intergeneric relationships can pose a particular problem as scholarship has long privileged intertextual relationships between works of the same genre. This paper treats the influence of Latin love elegy on Lucan's epic poem, *Bellum Civile*, by looking at two features of unmarked intertextuality: frequency and distribution. I use the Tesserae project to generate a dataset of potential intertexts between Lucan's epic and the elegies of Tibullus, Propertius, and Ovid, which are then aggregrated and mapped in Lucan's text. This study draws two conclusions: 1. measurement of intertextual frequency shows that the elegists contribute fewer intertexts than, for example, another epic poem (Virgil's *Aeneid*), though far more than the scholarly record on elegiac influence in Lucan would suggest; and 2. mapping the distribution of intertexts confirms previous scholarship on the influence of elegy on the *Bellum Civile* by showing concentrations of matches, for example, in Pompey and Cornelia's meeting before Pharsalus (5.722-815) or during the affair between Caesar and Cleopatra (10.53-106). By looking at both frequency and proportion, we can demonstrate systematically the generic enrichment of Lucan's *Bellum Civile* with respect to Latin love elegy.

Keywords
allusion; intertextuality; generic enrichment; Lucan; Latin epic; Latin love elegy; Tesserae

I INTRODUCTION

"There is no better way to penetrate the secrets of Lucan's workshop, to observe how the poem crystallized in his mind, than to examine passages where he borrows from, adopts, or echoes his predecessors." So wrote [Bruère, 1951 p. 222]. Bruère's contributions to the study of Lucan's allusive practice, especially his two articles co-authored with Thompson ([Thompson and Bruère, 1968; Thompson and Bruère, 1970]), brought a heightened awareness of the intertextual nature of the *Bellum Civile*, and in particular, its relationship to Virgil's *Aeneid*. The presence of intertextual influence from genres other than epic, however, has received far less attention.[1] Recent work by [Sannicandro, 2010; Caston, 2011; McCune, 2014] has sought to remedy this imbalance. These studies, however, have tended to

* This paper developed from dissertation research completed at Fordham University under advisor Matthew McGowan and an early version of this work was presented at the Institute for the Study of the Ancient World in the fall of 2015. I would like to thank Sebastian Heath, David Ratzan, Neil Coffee, James Gawley, Caitlin Diddams, and Neil Bernstein for their helpful feedback and would also like to thank the anonymous referees for their comments and suggestions.

[1] So, for example, [Groß, 2013 p. 279] writes in arguing for the intertextual influence of Horace on Lucan's poetry: "To date, the reception of Horace in Lucan has received sufficient attention neither in Lucan philology nor in Horatian scholarship. Because it was assumed that the reception of texts would only occur within the framework of their own genre, the Horatian odes and epodes were only seen as intertexts of lyrics, but never of epic poems. Accordingly, the analysis of Lucan's references to other poets was limited to other epics or tragedy, but lyrical predecessors were not considered." With respect to elegy specifically, note the comment of [Caston, 2011 p. 134], who writes that in looking at Lucan's "elegiac moments" she is able to "to highlight a genre that has been overlooked" in studies of Lucan's intertextuality.

emphasize localized readings, either referring to a limited number of elegiac source texts or treating a select group of episodes in Lucan's poem. In this paper, I use datasets drawn from the Tesserae Project at the University at Buffalo ([http1]) to systematically compare potential intertexts between Lucan's epic as a whole with reference to the complete works of the Latin love elegists Tibullus, Propertius, and Ovid.

II BACKGROUND

Systematic collections of elegiac references, no less analysis of these references, in epic poetry remain a desideratum in Latin literary criticism. Even within the epic genre, there are two works of traditional philological research which stand out for treating influence in a systematic and comprehensive manner, namely [Knauer, 1964] (with its subtitle "mit Listen der Homerzitate in der Aeneis") and [Nelis, 2001]. [Farrell, 2005 p. 107] has written that 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, having access to this kind of reference material would undoubtedly be useful. The existence of a book called "The *Bellum Civile* and Latin Love Elegy" would certainly appear in the bibliography of this study if it existed. [Coffee et al, 2012] remarks that traditional scholarly methods have avoided these kinds of comprehensive treatments of intertextuality because of the massive scholarly labor involved. Software is now available, however, to greatly reduce the procedural difficulty to which Coffee refers. Lists of potential intertexts can be compiled much more easily using Tesserae's web based tool ([http1]), which allows for the quick gathering of evidence for potential intertextuality between two texts, shifting scholarly labor from detection to analysis.

2.1 Problem of unmarked intertextuality

In his study of the intertextual relationship between Horace and Lucan, [Groß, 2013] observes that almost all intertextuality in classical poetry in unmarked, that is, it is not characterized by explicit signposts, but rather through implicit markers. Here, Groß follows the definition of unmarked intertextuality from [Helbig, 1996], who includes as two of the implicit markers 1. the "frequency" (*Frequenz*) of intertexts in the later text, and 2. the "distribution" (*Distribution*) of these intertexts, that is their location and relative density, throughout the text. This definition finds sympathy in two literary critical approaches to the problem of unmarked intertextuality, namely the "allusive system" discussed by [Farrell, 2005] and the "code model" discussed by [Conte, 1986]. Both Farrell and Conte argue that the relationship between two texts can be drawn to some degree by the volume of potential intertexts and their consistent presence throughout a target text. A collocation tool like Tesserae, by algorithmically determining and reporting a complete collection of correspondences, offers a formalization of Farrell's system and Conte's model. Moreover, the data collected from Tesserae results can be used to formalize Helbig's observation about frequency and distribution as implicit signposts for unmarked intertextuality. The analysis of Tesserae results can measure frequency by showing the number of times similarity in word use triggers a match and can measure distribution by showing which parts of the *Bellum Civile* show a greater or lesser number of matches.

2.2 Literature Review

In recent years, researchers at Tesserae have published a series of papers testing the assumptions of traditional Latin literary criticism against their algorithmic model ([Coffee et al, 2012; Coffee et al, 2013; Forstall et al, 2015]). These papers have used the first book of Lucan's *Bellum Civile* as their target text and Virgil's *Aeneid* as their source text, evaluating

the results of the automated tool against philological commentaries by assigning them, following [Thomas, 1986], values of "meaningful" and "not meaningful," as well as "interpretable" and "not interpretable." [Forstall et al, 2015] reports that scores assigned by the Tesserae algorithm correlate well with supervised assignments of meaning and interpretability.

The Tesserae publications have confirmed the traditional scholarly view that Lucan's poetic diction draws significantly on Virgil. That said, this research has consistently pointed the way towards wider applicability of algorithmically based methods for the study of intertextuality: [Coffee, 2012] suggests that systematic collection and measurement of textual similarities using a tool like Tesserae can build an "intertextual 'fingerprint'," that can be used to make meaningful comparisons between the poetic practices of different authors.

Important work on testing Tesserae search results is also being done by [Bernstein, 2013; Gervais, 2014; Bernstein, Gervais and Lin, 2015], who have concentrated on the platform's "macrophilological applications," that is ways in which the complete collection of search results for a given genre, author, or work can be used to draw conclusions, not about specific intertexts, but rather about larger patterns of intertextuality. [Bernstein, Gervais and Lin, 2015], in particular, in a study that looks at intertextual relationships in Latin hexameter poetry as a whole, argues that Tesserae can be used to generate an unlabeled dataset which captures the intertextual relationship between multiple Latin texts and can then be used as the basis for further analysis and interpretation.

III DATA

3.1 Texts

This study uses the following texts available from the Tesserae Github repository ([http2)].

The following editions of Latin epic poetry are used:

- Virgil, *Aeneid*: Greenough, J. B., ed. 1900. *Bucolics, Aeneid, and Georgics of Vergil*. Boston: Ginn & Co.
- Lucan, *Bellum Civile*: Weise, H., ed. 1835. *M. Annaei Lucani Pharsaliae Libri X*. Leipzig: G. Bassus.

The following editions of Latin love elegy are used:[2]

- Tibullus: Postgate, J.P., ed. 1915. *Tibulli aliorumque Carminum libri Tres.* Oxford: Oxford University Press.
- Propertius: Mueller, L., ed. 1898. *Sex. Propertii Elegiae*. Leipzig: Teubner.

[2] The collection listed above has been decided upon in order to align this work with that of Tesserae. It is obviously not the only arrangement available. [Pichon, 1902], for example, defined his sample as follows: the canonical works of Latin elegy mentioned in Ovid *Tristia* 4.10.53-54 and Quintilian *Institutiones* 10.1.93, to which he adds (or qualifies the inclusion of) Catullus, the *Corpus Tibullianum*, all of the *Heroides* regardless of authenticity, and certain poems from Ovid's *Tristia* and *Epistulae ex Ponto*. For Ovid, I use Ehwald's editorial decision to define the subset of Ovid's elegiac work which qualifies as erotic. Accordingly, the *Fasti*, *Ibis*, *Tristia* and *Epistulae ex Ponto* will not be used in this study. Along similar lines, I have based my decision to include the *Corpus Tibullianum* on Postgate's editorial decision and Tesserae's use of this edition.

- Ovid: Ehwald, R., ed. 1907. *Amores, Epistulae, Medicamina faciei femineae, Ars amatoria, Remedia amoris.* Leipzig: Teubner.

I have used the volumes listed above so that meaningful comparisons can be made with Tesserae studies which have already been published as well as those being conducted by other researchers. This follows the recommendation of [McGillivray, 2014], who argues that it is methodologically critical to work within a "collaborative research paradigm," that is to work from a common set of texts and to build directly upon existing tools and frameworks in an effort to maintain replicability in literary research. I have published the data set and the code used to generate the tables and figures on Github [http3].

3.2 Tesserae Search Results

Tesserae describes itself as a framework for "detecting allusions" in Latin poetry. More precisely, it is a search tool designed: 1. to compare the texts of two authors by looking for shared words, and 2. to return a list of similar passages scored for significance on a scale from 2 to 10. All results require that the units of text under consideration contain a minimum of two shared words. Once this requirement is met, matches are scored algorithmically based on two factors: word frequency and phrase density. Descriptions of the scoring algorithm can be found in [Forstall et al, 2015 p. 504].[3] Word frequency refers to how common or uncommon a matched word is within the two texts; phrase density refers to the number of interstitial words separating the matched words. These parameters are designed to make explicit the formal criteria that scholars have traditionally applied implicitly when identifying an allusion. Accordingly, less common words which are adjacent receive higher scores than more common words which are separated by gap of several words. For example, the adjacent collocation of the rare words *livor edax* (Ovid *Amores* 1.15.1: *Quid mihi Livor edax, ignavos obicis annos…* ~ Lucan *Bellum Civile* 1.288: *Livor edax tibi cuncta negat: gentesque subactas*) receives a score of 10, while a separated collocation of the very common words *quod* and *te* (e.g., Ov. *Amores* 2.9b.47 ~ Lucan *Bellum Civile* 9.854) receives a score of 3.

On the one hand, by its nature, the Tesserae algorithm collects matches in an unrestricted manner. That is, it returns as many matches as fit its criteria—the text analytic equivalent of trawl fishing or strip mining. The result is a high number of false positives or dubious, semantically empty connections, especially due to ambiguity in lemmatization ([Forstall et al, 2015]). These matches correlating with the low end of the Tesserae scoring scale will be largely ignored in this study. On the other hand, [Coffee et al, 2012; Forstall et al, 2015] have shown that the high scores (that is, scores 10 and 9) generated by the Tesserae scoring algorithm correlate with meaningful and interpretable results, and that the next tier of scores (scores 8 and 7) correlate with meaningful results. Accordingly, the high scores will be the focus of this study.

[3] [Forstall et al., 2015] report the equation for the Tesserae scoring system as follows:

$$\text{score} = \ln\left(\frac{\Sigma\frac{1}{f(t)} + \Sigma\frac{1}{f(s)}}{d_t + d_s}\right)$$

where "$f(t)$ is the frequency of each matching term in the target phrase; $f(s)$ is the frequency of each matching term in the source phrase; d_t is the distance in the target; d_s is the distance in the source."

The data used for the study of intertextual frequency and distribution was gathered using version 3 of the Tesserae search interface.[4] The following parameters were used for these searches:

- unit = line
- feature = stem
- stopsize = 20
- stbasis = corpus
- stopwords = qui quis sum et in is hic non ego ut cum tu ad ille quod ab si atque neque sed
- max_dist = 10
- dibasis = freq
- cutoff = 0

Full definitions of the terms used in this list are available at the Tesserae site [http4].[5]

Defaults were used where possible to ensure to the greatest degree possible comparability between this study and other Tesserae-based studies. One exception is "line" as the unit for this study; the maps of intertextual distribution use line numbers for the x-axis and the number of matches per line for the y-axis.

Tesserae results yield the following information:

- Result number
- Target Location[6]
- Target Text
- Source Location
- Source Text
- Shared words
- Score

Here is an example of a record from the .csv file returned by a Tesserae search between the *Bellum Civile* and Ovid's *Amores*:

---

[4] The Tesserae searches used in this study were run between September 13-16, 2015.

[5] The following comments should be helpful in getting a quick understanding of the Tesserae settings. Note that these are the settings as reported in the "Tesserae V3 results" when exported as csv, txt, or xml. "Feature" determines how words are treated by the algorithm for making comparisions; under the setting "stem" (called "Lemma" in the Tesserae interface), the program "returns sets of parallels between texts where the matched words have the same dictionary headwords," that is, *amor* in one text matches *amoris*, *amori*, *amorem*, etc. in another. "Stopwords" is the default Tesserae stoplist, that is, the list of words ignored in this study; the basis for determining the stopword list ("stbasis") is the corpus of all works available in Tesserae. Accordingly, the algorithm ignores the 20 most commonly appearing words in this corpus. "Max_dist" refers to the maximum distance that the algorithm uses for its window for matches; that is words in a text must be between two and ten words apart from each other, counting inclusively, to yield a result. "Dibasis" refers to the manner in which the algorithm accounts for distance between words; according to [http4], with the setting "freq," Tesserae "attempts to zero in on the most relevant words in an allusion, measuring the distance only between the phrase's two most infrequent words." "Cutoff" refers to the minimum score returned by Tesserae; I have set it at zero to gather the full range of Tesserae results, although scores below 7 will be dropped for most parts of this study as noted below.

[6] The terms source text and target text are used in the study of intertextuality, in Latin literature and elsewhere, to refer to the relationship of texts in the "traditional citation of parallel passages," ([Fowler 1997, p. 14]) most commonly with a passage in an earlier text (the source) corresponding to a passage in a later text (the target).

RESULT	TARGET_LOC	TARGET_TXT	SOURCE_LOC	SOURCE_TXT	SHARED	SCORE
11	"luc. 1.288"	"**Livor** **edax** tibi cuncta negat: gentesque subactas"	"ov. am. 1.15.1"	"Quid mihi **Livor** **edax**, ignavos obicis annos,"	"edax; liuor"	10

IV METHODS

Following [Helbig, 1996], I have designed this study to measure two of his criteria for unmarked intertextuality: frequency and distribution.

Frequency is measured by aggregating the count of matches by author or work. Tesserae returns search results as a .csv and these results are then converted into dataframes for processing with the Python Pandas module ([McKinney, 2012]). Raw counts are normalized to matches per 100 lines to account for differing text lengths. Lengths were determined using the Tesserae texts given in the section "2.1 Texts" above. Tesserae scores below 7 are discarded to reduce "noise" and to restrict the analysis to scores which have been shown to yield the most significant results ([Forstall et al, 2015; Bernstein, Gervais and Lin, 2015]).

Distribution is measured by using Pandas to aggregate the total number of Tesserae matches from the elegists above score 7 for each line in the *Bellum Civile*. Lines from the *Bellum Civile* with no matches from Tesserae are assigned a count of zero. Because the goal of studying the distribution of unmarked intertextuality is to see how the feature appears throughout the entire work, rather than in any given line, the counts of matches per line are smoothed by taking the running average of scores within a specific window, here 25 lines. This reduces the line-by-line variability in counts while providing a better sense of potential intertextual density in different sections of the target text. The smoothed scores are then mapped by book using line numbers for the x-axis and the smoothed counts of Tesserae matches for the y-axis.

V DISCUSSION

5.1 Measuring Frequency

Table 1 shows the total number of Tesserae matches with Lucan's *Bellum Civile* as the target text for each of the elegists and their works, with Virgil's *Aeneid* included for comparison. The raw counts, however, are not sufficient to compare the authors/works, because they are of greatly varying lengths. For example, Propertius's four books of elegies have a total of 3,982 lines compared to the 9,896 lines of Virgil's *Aeneid*. Accordingly, it is necessary to normalize these scores so that they can be compared more usefully. By normalizing the counts of Tesserae results to matches per 100 lines, we get a different picture than the received view of intertextuality in Lucan, as shown in Table 2. The number of matches between Lucan and Virgil is much higher that the number between Lucan and any of the elegists. When normalized, however, the elegists show on average only a 17% difference, and in the case of Ovid's *Heroides*, the matches per 100 lines is slightly higher (469.67 in the *Heroides* versus 468.72 in the *Aeneid*).

Figure 1 shows the total number of matches by score for Virgil. As noted above, research on Tesserae results shows that results scoring 7 or above correlate best with meaningful results. Accordingly, in order to reduce the signal-to-noise ratio and limit the remainder of this part of the study to only the most meaningful results, scores below 7 will not be considered below. Figure 2 shows the total number of matches by score above this threshold for Virgil.

By combining the strategies listed above of normalizing the scores and concentrating on the scores above a certain threshold, we can now compare the elegists to each other, again with the *Aeneid* included as a baseline. Looking at Figure 3, we can see that the shared diction between Virgil and Lucan as measured by the Tesserae algorithm, that is with consideration of word frequency and phrase density, is higher than for any of the elegists. This is hardly surprising considering that the history of scholarship on Lucan has privileged this relationship far beyond that of any other poetic predecessor. Rather the conclusion that can be drawn from this chart is that the minimal scholarly attention that has been paid to the influence of the elegists on Lucan, based on the Tesserae data, is out of proportion with the attention paid to the *Aeneid*.

The score most likely to be meaningful and interpretable, that is score 10, is exceptionally rare in all authors, consistently appearing less than one time per 100 lines. For score 9, the elegists show roughly half as many matches per 100 lines (average 2.77 per 100 lines; 48.7% of Virgil's 5.70). The count for Ovid's *Heroides* is notable for standing out as being much higher than the others (4.25 matches per 100 lines; 90.4% of Virgil's 5.70), suggesting that the density of specific allusions noted between these poems and the *Bellum Civile* by previous scholars (e.g. [Bruère, 1951; Sannicandro, 2010]) is supported by the evidence.

We find an example of a meaningful and interpretable match at line 7.590 of the *Bellum Civile*, when the poet addresses Brutus on the battlefield at Pharsalus, encouraging him not to enter the fray and kill Julius Caesar: "*Ne rue per medios nimium temerarius hostis*, Do not rush too rashly out through enemy lines." The collocation *nimium temerarius* appears in only two other passages of classical Latin literature, both in love elegy. Propertius uses it at 2.8.13 to chastise himself for being "too rash" in supporting an unfaithful lover, and Ovid uses the same language at *Ars amatoria* 2.83 to describe Icarus's insolence in disobeying his father. Lucan adopts elegiac diction at a critical moment to undermine Brutus's potential for epic glory.[7] The rarity of this collocation and the fact that the words are adjacent within their lines lead to Tesserae scores of 9 between the *Bellum Civile* and both elegiac works, and so here we see the correlation of high scores between the elegists and Lucan with respect to lines from the former which are, to use the description of [Hinds, 2007 p. 119], "thematically grounded" in the latter.

With the meaningful and non-interpretable scores (8 and 7), that is those which contribute to the general elegiac texture of the work and show evidence of elegy as a code model for the *Bellum Civile*, we see a similar pattern. Virgil's *Aeneid* again shows a higher number of matches per 100 lines, but not by the overwhelming margin that scholarly attention between the two genres would suggest: for score 8, the elegists show a little more than two-thirds as many matches per 100 lines (average 14.95 per 100 lines; 70.0% of Virgil's 21.38), and for score 7, more than three-quarters as many matches (average 68.0 per 100 lines; 77.5% of Virgil's 87.76).

5.2 Mapping Distribution

In addition to data about frequency, Tesserae searches for the elegists and the *Bellum Civile* also provide us with data about the location of the matches. Figure 4 shows a plot of the

[7] As [Gorman, 2001 p. 270] writes: "In a most anti-epic stance, Lucan actually admonishes Brutus to avoid the fight, in direct contravention to what a hero ought to be doing. Courage, the defining quality of the epic warrior, has become the fault of rashness (*nimium temerarius*)."

distribution of aggregated Tesserae matches per line for the complete set of elegiac texts as the source text and book 1 of the *Bellum Civile* as the target text. We learn from this text map that the matches are randomly distributed throughout the book, suggesting that elegy is functioning as a code model for the *Bellum Civile*. The line with the highest number of matches (15) is Lucan 1.61 (*...inque vicem gens omnis amet: Pax missa...*), likely reflecting the presence of "love" (*amet*) in this line, the central theme of the elegiac poems under consideration; the average number of matches per line is 1.69, and 65% (449 out of 695) of lines in book 1 have at least one match.

That said, the elegiac weight of any given line gives us information that is perhaps too localized. In order to get a better sense of which episodes show a sustained interaction with elegy, it is preferable to plot the running average of matches within a certain window of lines. Figure 5 shows a plot of this running average for each book of the *Bellum Civile* using a window of 25 lines.

As compared to the line-by-line mapping of counts from the Tesserae results, plotting by window confirms once again that matches are distributed consistently throughout the poem. It also, and much more usefully as a prompt for further investigation, provides a clearer idea of where in the *Bellum Civile* additional research into elegiac influence would be most fruitful. Certain peaks in these text maps corroborate the work of previous scholarship on elegiac influence in Lucan. For example, we see pronounced upticks in the average number of matches at the end of book 5 for Pompey and Cornelia's meeting before Pharsalus (Lucan 5.722-815; see [Sannicandro, 2010; Bruère, 1951]), and at the beginning of book 10 for the appearance of Cleopatra (Lucan 10.53-106, 172-192; see [McCune, 2014; Groß, 2013]). Another interesting takeaway from this view of the distribution of elegiac language in the *Bellum Civile*, is the presence of peaks in parts of the epic where scholarly research into the interaction of these two genres has not been previously focused, as, for example, in the book 1 proem (1.1-32) or Cato's speech to his troops in book 9 (9.222-283).

VI Conclusion

The automated detection and measurement of intertextuality, in particular the definition of "allusion" formalized by Tesserae, offers the insight into the "poet's workshop" that Bruère described. It allows researchers to systematize and quantify the intertextual readings brought out in traditional, qualitative literary analysis of Latin poetry. This study uses the evidence generated by Tesserae to support Helbig's conception of frequency and distribution as implicit indicators of unmarked intertextuality. It also provides data to the idea that an "allusive system" can be deduced from a mass of textual similarities. The large number of scores 7 and higher for all three Latin love elegists, and the fact that these results are not confined to a small number of locations in the text suggest that this genre acts to some degree as a "code model" for the *Bellum Civile*.

[Coffee et al, 2013 p. 227] showed how Tesserae could be used to corroborate the dominant scholarly opinion about Lucan's reliance on Virgil "for the basic idiom of epic." What is less in line with the traditional view of Lucan's allusive practice however is to what extent the results given above demonstrates elegy's contribution to his epic "idiom." The frequency of intertextual correspondence may be less than for Virgil's *Aeneid* and their distribution more diffuse, but this study was not meant to disprove Virgil's authority or deny his influence on the later epic poet. Rather it is meant to show that another genre can compete for some space in Lucan's attention to poetic predecessors. Scholarly consideration of the effect of elegy on the *Bellum Civile* is miniscule compared to the massive amount of research on Lucan's poetic

relationship to, or even dependence, on Virgil as a poetic model. Yet, as we see here, the gap in the evidence is not as wide as the record of scholarship suggests.

This distant reading approach to verifying the allusive system and the code model is a starting point. [Bernstein, Gervais and Lin, 2015] comments that a limitation of the Tesserae platform is that it does not yet offer the "sensitive assessment of significance" that is taken for granted in traditional forms of Latin literary criticism. Along these lines, the test for frequency in this study does not address whether individual results are in fact meaningful or interpretable, and the test for distribution, while pointing out episodes worthy of further scrutiny, has not evaluted them. [Forstall et al, 2015] reminds users that Tesserae data must still be interpreted within the established scholarly conversation concerning intertextuality in Latin poetry. Examination of the peaks and troughs of the text maps of allusive distribution forms an excellent starting point for a more "sensitive assessment" of the intertextual relationship between Lucan and the love elegists in the future and also provides a more empirical method for investigating the "generic enrichment" ([Harrison, 2007; Harrison, 2013]) found in Latin poetry.

VII TABLES AND FIGURES

7.1 Tables

	Work	Tesserae Matches
Tibullus	*Elegies*	6,860
Propertius	*Elegies*	13,812
Ovid	*Amores*	9,966
	Ars Amatoria	9,235
	Heroides	18,646
	Medicamina	285
	Remedia Amoris	3,711
Virgil	*Aeneid*	46,385

Table 1. Total number of Tesserae matches for Lucan's *Bellum Civile* as the target text and the Latin love elegists and Virgil's *Aeneid* as source texts.

	Work	Lines	Tesserae Matches	Count per 100 Lines
Tibullus	*Elegies*	1,929	6,860	355.62
Propertius	*Elegies*	3,982	13,812	346.86
Ovid	*Amores*	2,445	9,966	407.61
	Ars Amatoria	2,330	9,235	396.35
	Heroides	3,970	18,646	469.67
	Medicamina	100	285	285.00
	Remedia Amoris	814	3,711	455.90
Virgil	*Aeneid*	9,896	46,385	468.72

Table 2. Number of Tesserae matches normalized per 100 lines for Lucan's *Bellum Civile* as the target text and the Latin love elegists and Virgil's *Aeneid* as source texts.

	Work	Score 10 (per 100)	Score 9 (per 100)	Score 8 (per 100)	Score 7 (per 100)
Tibullus	*Elegies*	0.31	2.64	18.56	76.31
Propertius	*Elegies*	0.33	3.16	15.77	67.88
Ovid	*Amores*	0.49	2.74	18.45	75.83
	Ars Amatoria	0.13	2.79	14.98	72.27
	Heroides	0.13	4.26	18.49	79.07
	Medicamina	1.00	2.00	6.00	33.00
	Remedia Amoris	0.25	1.84	12.41	71.62
Virgil	*Aeneid*	0.87	5.70	21.38	87.76

Table 3. Number of Tesserae matches by score (at or above a threshold of 7) normalized per 100 lines for Lucan's *Bellum Civile* as the target text and the Latin love elegists and Virgil's *Aeneid* as source texts.

7.2 Figures

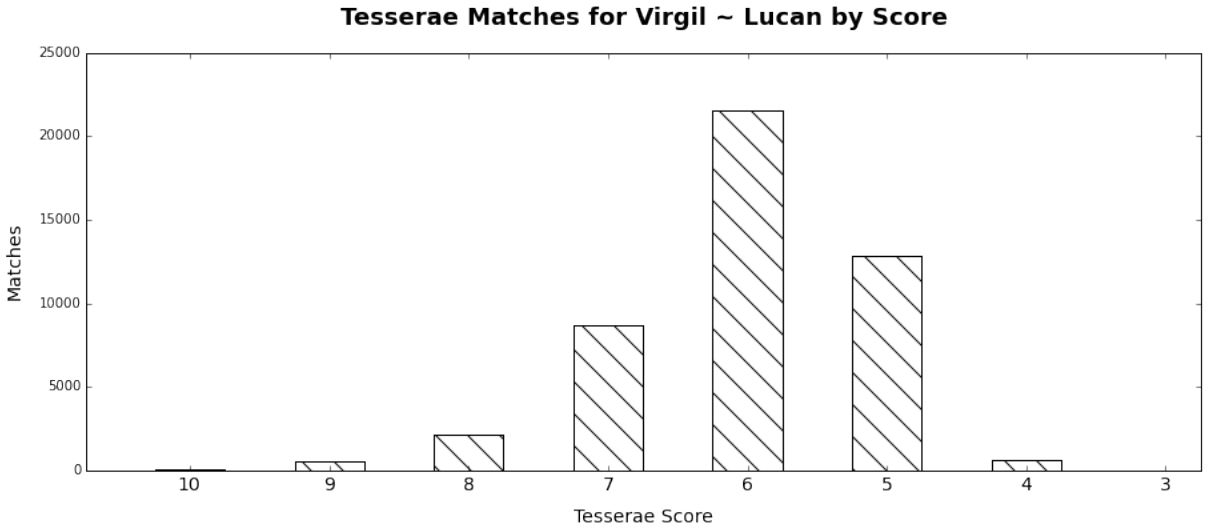


Figure 1. Number of Tesserae matches by score for Virgil's *Aeneid*.

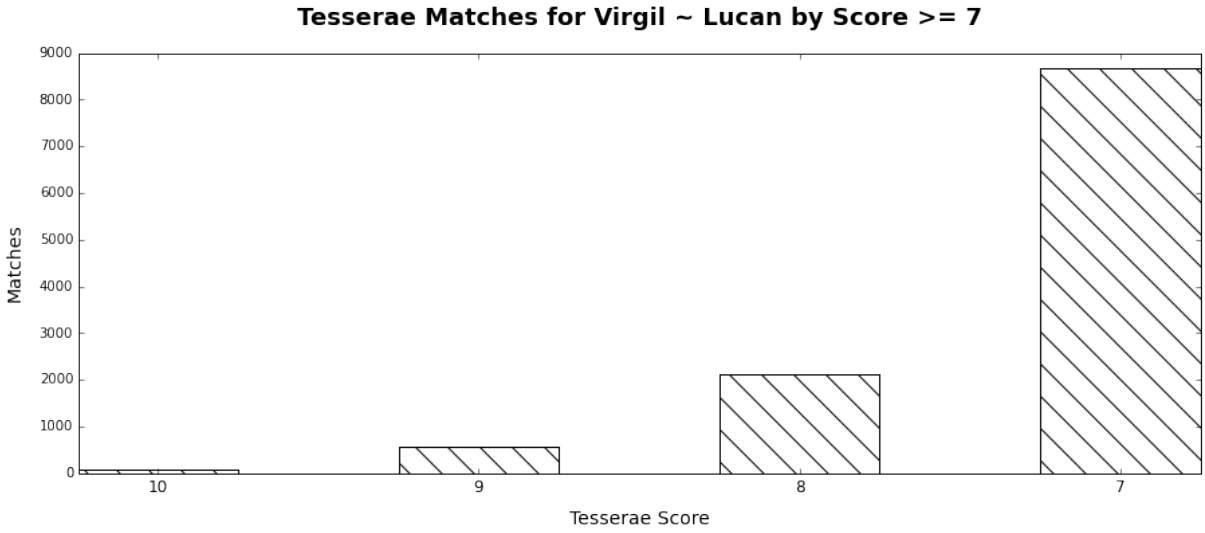


Figure 2. Number of Tesserae matches by score for Virgil's *Aeneid* at or above a threshold of 7.

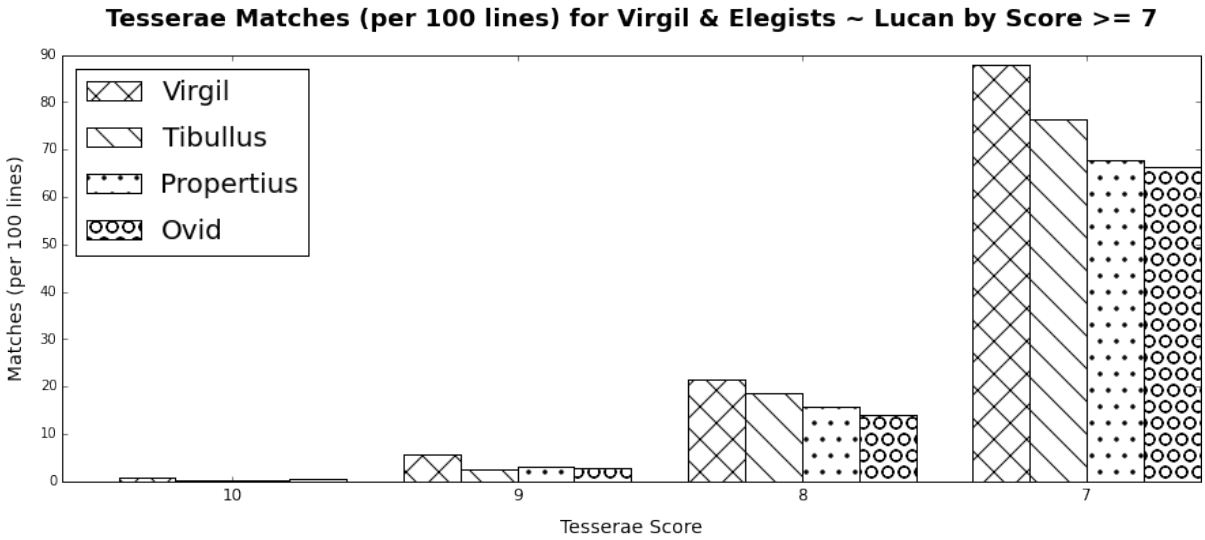


Figure 3. Chart showing the number of Tesserae matches by score (at or above a threshold of 7) for the Latin love elegists and Virgil's Aeneid normalized to count per 100 lines as noted in Table 3.

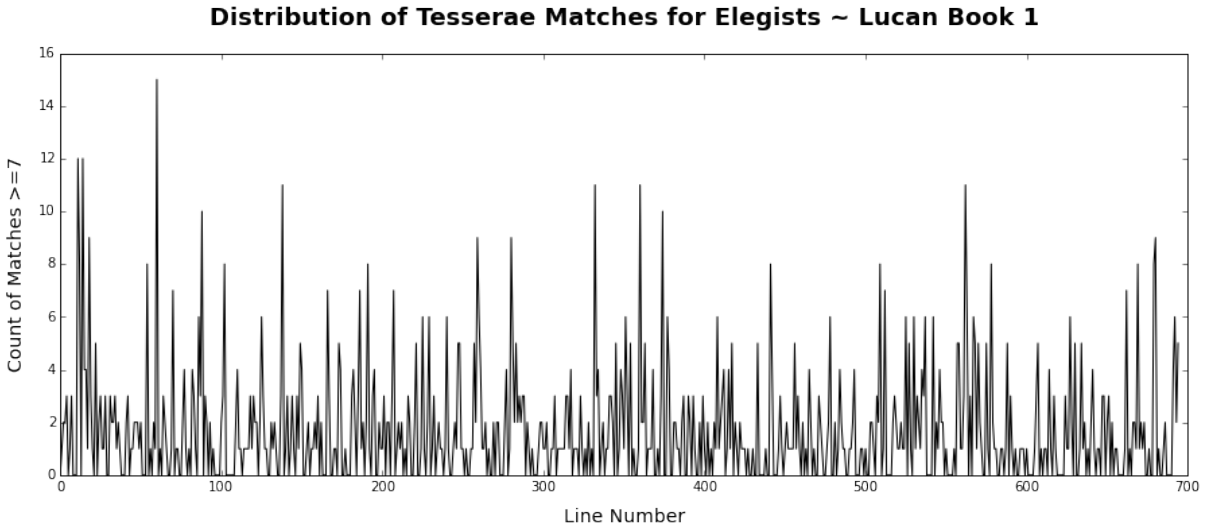
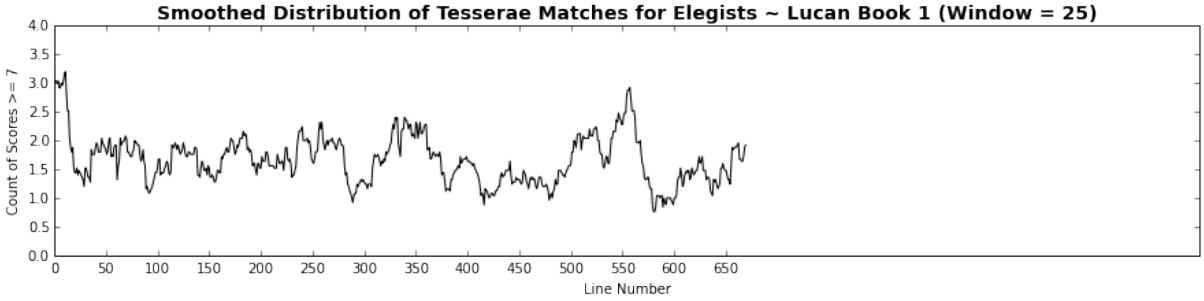
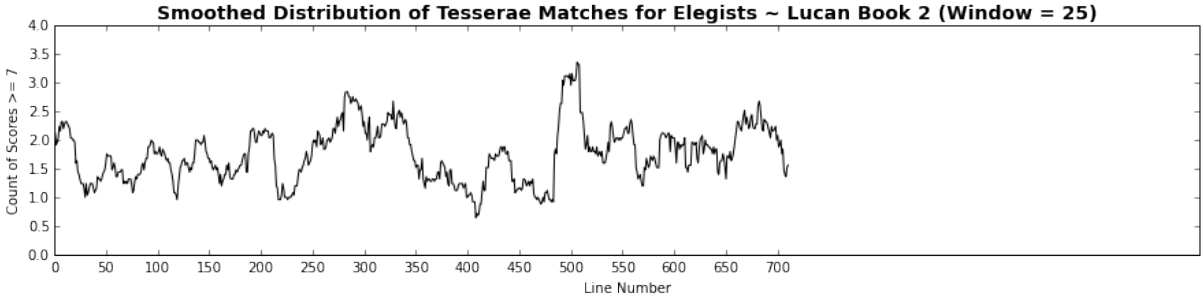
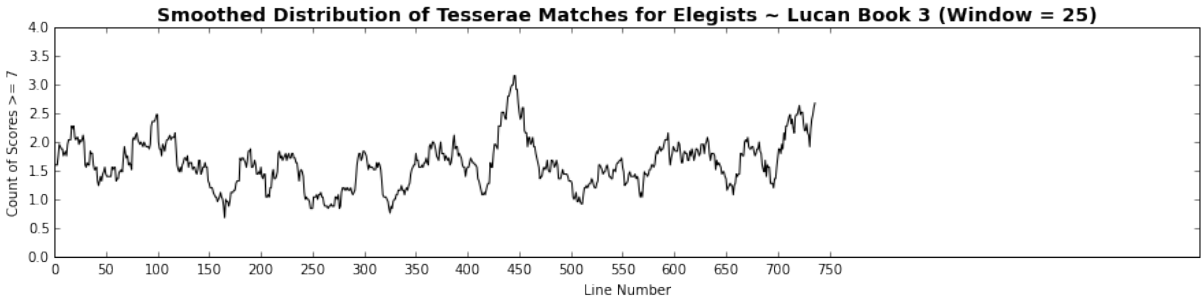
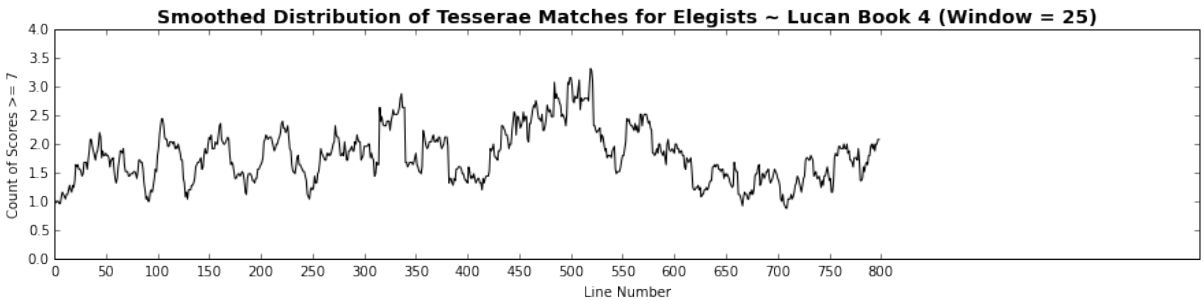


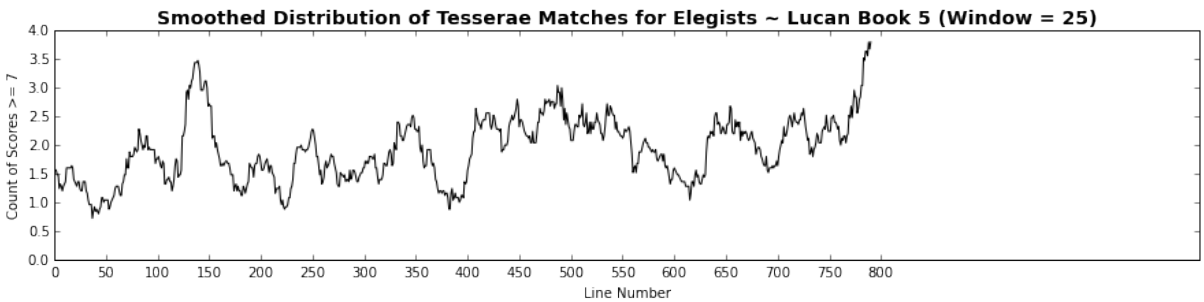
Figure 4. Total count of Tesserae matches per line of book 1 of Lucan's *Bellum* Civile for the complete set of elegiac texts.

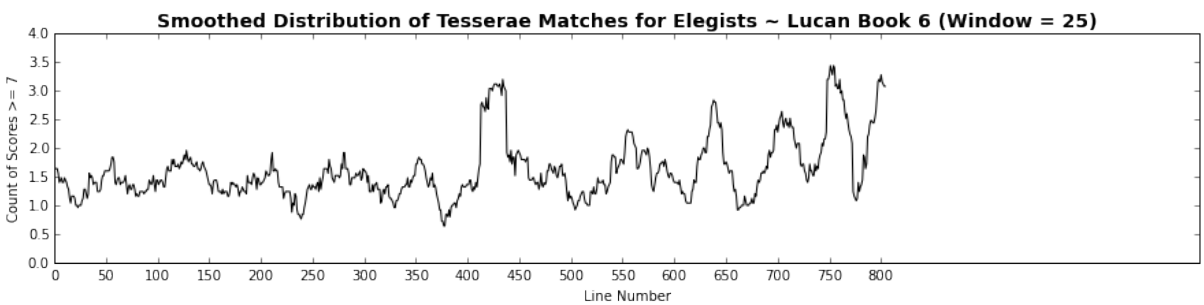

Smoothed Distribution of Tesserae Matches for Elegists ~ Lucan Book 1 (Window = 25)
Count of Scores >= 7
Line Number


Smoothed Distribution of Tesserae Matches for Elegists ~ Lucan Book 2 (Window = 25)
Count of Scores >= 7
Line Number


Smoothed Distribution of Tesserae Matches for Elegists ~ Lucan Book 3 (Window = 25)
Count of Scores >= 7
Line Number


Smoothed Distribution of Tesserae Matches for Elegists ~ Lucan Book 4 (Window = 25)
Count of Scores >= 7
Line Number


Smoothed Distribution of Tesserae Matches for Elegists ~ Lucan Book 5 (Window = 25)
Count of Scores >= 7
Line Number


Smoothed Distribution of Tesserae Matches for Elegists ~ Lucan Book 6 (Window = 25)
Count of Scores >= 7
Line Number

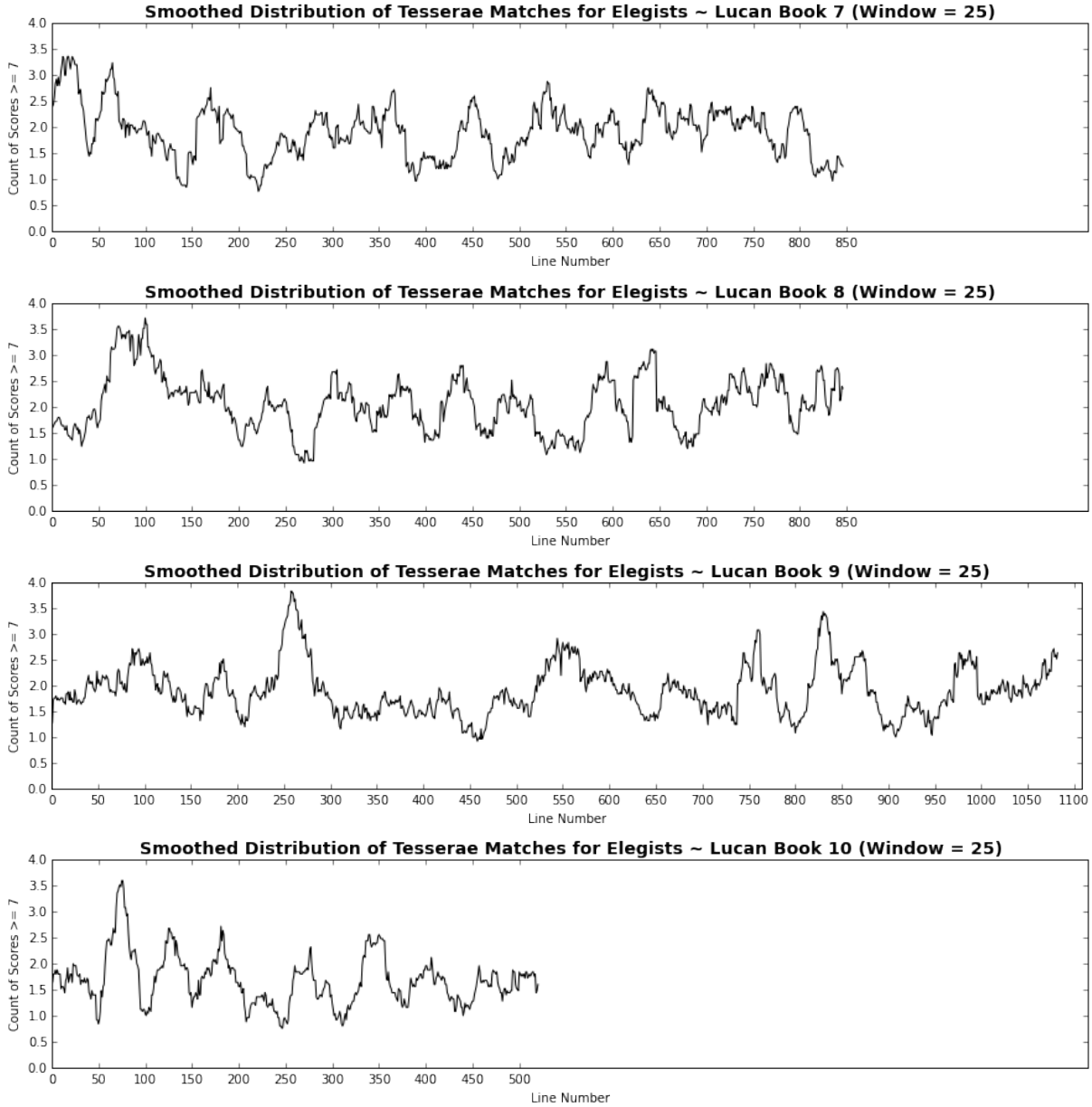


Figure 5. Average count of Tesserae matches for the elegists in the ten books of the *Bellum Civile*, smoothed using a window of 25 lines.

References

Bernstein N. Using Tesserae to extend the philological commentary. Presented at *Word, Space, Time: Digital Perspectives on the Classical World*, 2013. Buffalo, NY. 5-6 April.

Bernstein N, Gervais K and Lin W. Comparative rates of text reuse in classical Latin hexameter poetry. 2015;9(3). Available from: http://www.digitalhumanities.org/dhq/vol/9/3/000237/000237.html. Accessed 2016-02-27.

Bruère R.T. Lucan's Cornelia. *Classical Philology.* 1951;46(4):221-236.

Caston R.R. Lucan's elegiac moments. In Asso P., ed. *Brill's Companion to Lucan*. Leiden: Brill; 2011: 133-152.

Coffee N., Gawley J., Forstall C.W., Scheirer W.J., Johnson D., Corso J.J. and Parks B. Modelling the interpretation of literary allusion with machine learning techniques. *Journal of Digital Humanities* 2014;3(1): http://journalofdigitalhumanities.org/3-1/modelling-the-interpretation-of-literary-allusion-with-machine-learning-techniques/. Accessed 2016-03-01.

Coffee N., Koenig J.-P., Poornima S., Ossewaarde R., Forstall C.W. and Jacobson S. Intertextuality in the digital age. *Transactions of the American Philological Association.* 2012;142(2):383-422.

Conte G.B. *The Rhetoric of Imitation: Genre and Poetic Memory in Virgil and Other Latin Poets.* Cornell University Press (Ithaca, NY), 1986.

Farrell J. Intention and intertext. *Phoenix.* 2005;59(1):98-111.

Forstall C.W., Coffee N., Buck T., Roache K. and Jacobson S. Modeling the scholars: Detecting intertextuality through enhanced word-level N-gram matching. *Digital Scholarship in the Humanities.* 2015;30(4):503-515.

Fowler D. On the shoulders of giants: intertextuality and classical studies. *Materiali e discussion per l'analisi dei testi classici*. 1997; 39: 13-34.
Gervais, K. Flavian intertextuality: A digital approach. Presented at 35th Conference of the Australasian Society for Classical Studies, 2014. Aukland, NZ: 29 January.
Gorman, V.B. Lucan's epic *Aristeia* and the hero of the *Bellum Civile*. *Classical Journal*. 2001;96(3): 263-290.
Groß D. *Plenus Litteris Lucanus: Zur Rezeption der Horazischen Oden Und Epoden in Lucans Bellum Civile.* Verlag Marie Leidorf (Rahden/Westfalen), 2013.
Harrison S.J. *Generic Enrichment in Vergil and Horace.* Oxford University Press (Oxford), 2007.
Harrison S.J. Introduction. In Papanghelis T.D., Harrison S.J., Frangoulidis S., eds. *Generic Interfaces in Latin Literature: Encounters, Interactions and Transformations*. Berlin: De Gruyter; 2013: 1-15.
Helbig J. *Intertextualität und Markierung.* Universitätsverlag C. Winter (Heidelberg), 1996.
http1 http://tesserae.caset.buffalo.edu
http2 https://github.com/tesserae/tesserae/tree/master/texts/la
http3 https://github.com/diyclassics/measuring-allusion
http4 http://tesserae.caset.buffalo.edu/help_advanced.php
Hinds, S. Martial's Ovid/Ovid's Martial. *Journal of Roman Studies*. 2007; 97(1): 113-154.
Hübner U. Episches und Elegisches am Anfang des dritten Buches der *Pharsalia. Hermes.* 1984;112(2):227-239.
Knauer G.N. *Die Aeneis und Homer: Studien Zur Poetischen Technik Vergils, Mit Listen Der Homerzitate in Der Aeneis.* Vandenhoeck & Ruprecht (Göttingen), 1964.
McCune B.C. Lucan's *Militia amoris*: Elegiac expectations in the *Bellum civile. Classical Journal.* 2014, Jan;109(2):171-198.
McGillivray B. *Methods in Latin Computational Linguistics.* Brill (Leiden), 2014.
McKinney W. *Python for Data Analysis Data Wrangling with Pandas, NumPy, and IPython.* O'Reilly Media (Sebastopol, CA), 2012.
Nelis D.P. *Vergil's Aeneid and the Argonautica of Apollonius Rhodius.* Francis Cairns (Leeds), 2001.
Pichon R. *De sermone amatorio apud Latinos elegiarum scriptores.* Hachette (Paris), 1902.
Sannicandro L. *I Personaggi Femminili Del Bellum Civile Di Lucano.* Marie Leidorf (Rahden), 2010.
Thomas R.F. Virgil's *Georgics* and the art of reference. *Harvard Studies in Classical Philology.* 1986;90:171-198.
Thompson L. and Bruère R.T. Lucan's use of Virgilian reminiscence. *Classical Philology.* 1968;63(1):1-21.
Thompson L. and Bruère R.T. The Virgilian background of Lucan's fourth book. *Classical Philology.* 1970;65(3):152-172.