

Motivation



We are interested in understanding the nature of the sound that is constrained by elegiac couplets – does it reflect the voice of the poet, or the general style of the elegiac form?

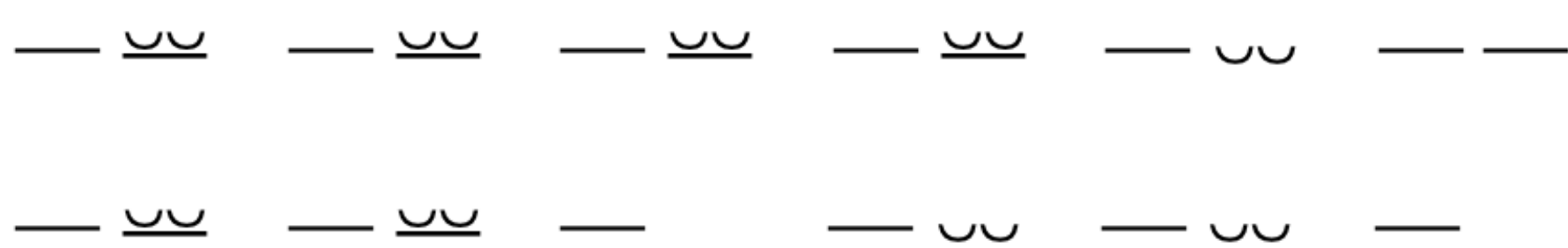
Within the Digital Humanities, stylistic studies have been produced for a wide variety of literature, including poetry. Existing feature sets and analysis techniques have most often examined texts at the word-level. A word-level examination captures only part of the underlying sound content of a poem, which is fundamental to its composition. Here we introduce a variety of sound-based statistical features found to be useful descriptors of Latin poetics.

In this work, we look at the role repetitive sound plays in the Latin elegiac couplet, where just a single character-level bi-gram can be a defining component of the form¹. We are working to incorporate our feature sets and classification components into the University at Buffalo's *Tesserae*² project, an online tool which provides scholars studying Latin poetry easy access to sophisticated textual analysis tools.

This work is part of an ongoing study^{1,3} of repetitive sound and its relationship to style in poetry.

Elegiac Couplets

The elegiac meter⁴ is used for a variety of themes, most notably Love⁵. The elegiac couplet is a pair of two different one-line "verses":



In the above, — represents a long syllable and ~ a pair of short syllables; the two symbols superimposed represent the poet's choice of either one long or two shorts. The first verse is identical to a verse of dactylic hexameter; the second, often called the "pentameter" verse of the couplet, is shorter by two half-feet.

Catullus 85

ōd' ēt āmō. quār' īd faciām, fōrtāssē rēquīrīs.
nēscīō, sēd fīērī sēntī' ēt ēxcrūciōr.

I hate and I love. Perhaps you ask why I do it?
I don't know, but I feel it happening,
and I am in torment.



Tibullus 1.5 – lines 75 & 76

nēscīō quīd fūrtīvūs āmōr pārāt. ūtērē, quāēsō,
dūm līcēt : īn līquīdā nāt tībī līntēr āquā.

Sneaky Love is up to something.
Enjoy it while you can, I beg:
your boat sails in clear waters.

Ovid Amores 1.10 – lines 29 & 30

sōlā vīrō mūliēr spōliīs ēxultāt ādēmpītis,
sōlā lōcāt nōctēs, sōlā līcēndā vēnīt,

Alone woman delights
in what she steals from a man,
Alone she hires out her nights,
alone she comes up for sale.

The Functional n-gram Analysis

Observation: Sound plays a fundamental role in an author's style, particularly for poets.

The functional n-gram is a feature for stylistic analysis, whereby the power of the Zipfian distribution is realized by selecting the n-grams that occur most frequently as features, while preserving their relative probabilities as the actual feature element.

Feature: The Functional n-gram

$$P(e_n | e_{n-N+1}^{n-1}) = \frac{C(e_{n-N+1}^{n-1} e_n)}{C(e_{n-N+1}^{n-1})} \text{ iff } \text{freq}(e_{n-N+1}^{n-1} e_n) > \phi$$

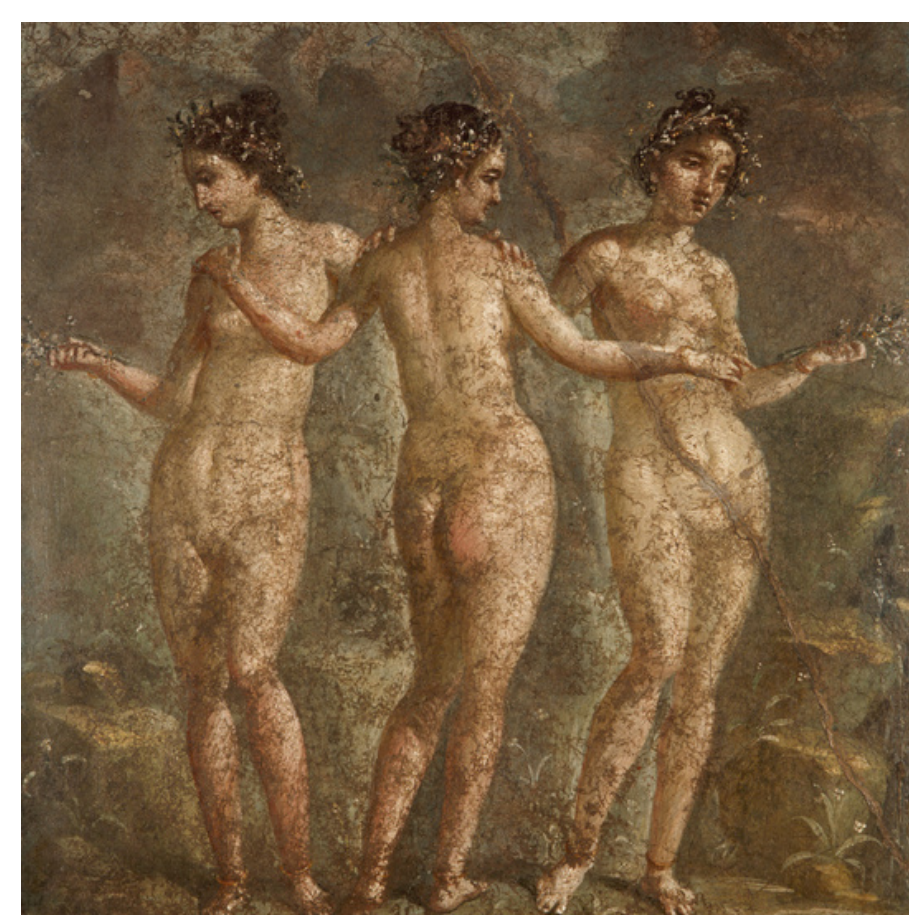
In this work, we consider primitive sound elements as functional character level bi-grams.

Latin Elegists considered in this study:

Catullus
Ovid
Propertius
Tibullus

Other Latin poets considered in this study:

Horace
Juvenal
Lucan
Lucretius
Statius
Vergil

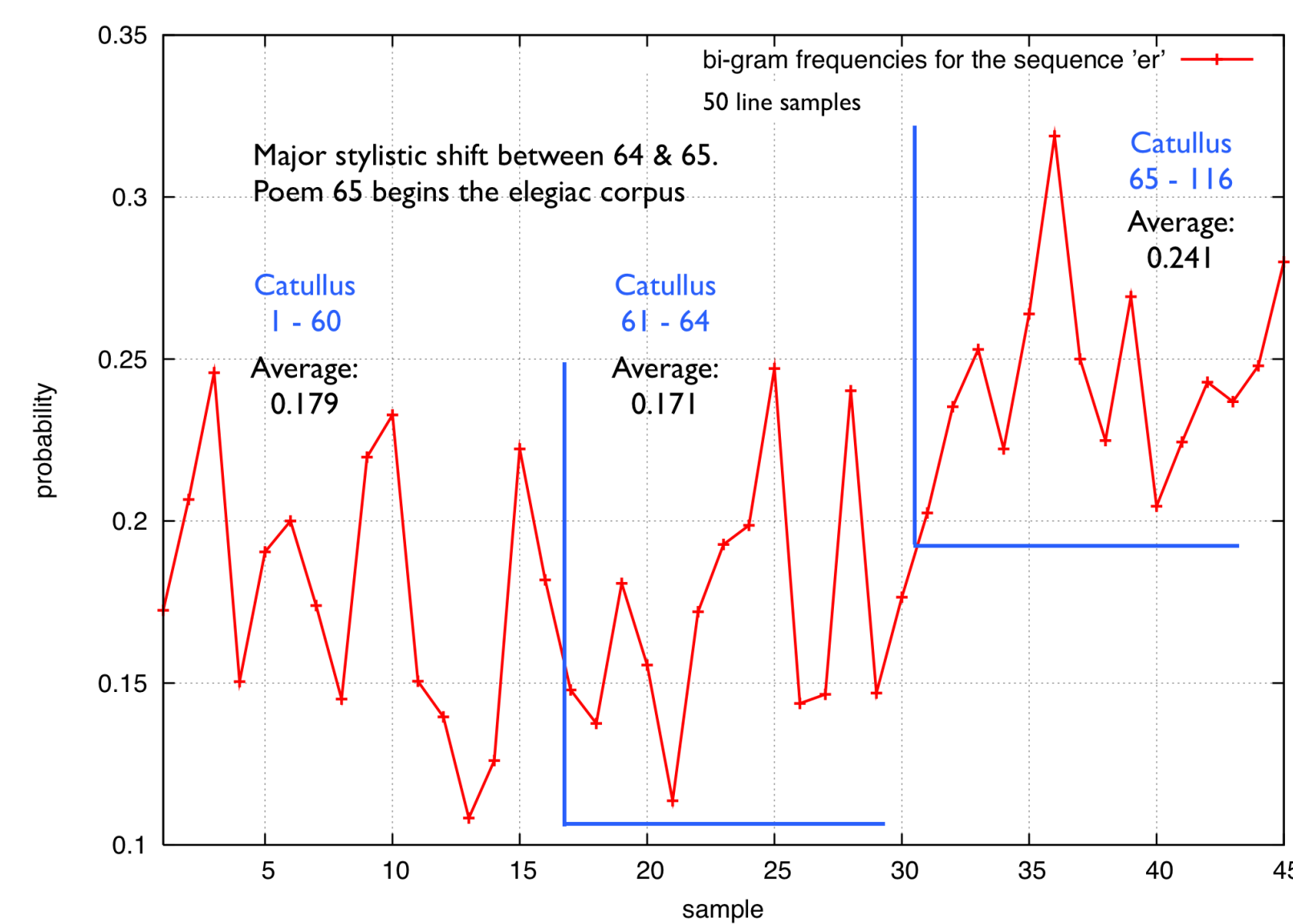


Functional n-grams for elegiac couplets:

er – top bi-gram that is common to all poets considered
nt – bi-gram with the greatest metrical variation
um – bi-gram sensitive to meter signal
am – bi-gram sensitive to meter signal

The Significance of the bi-gram *er*

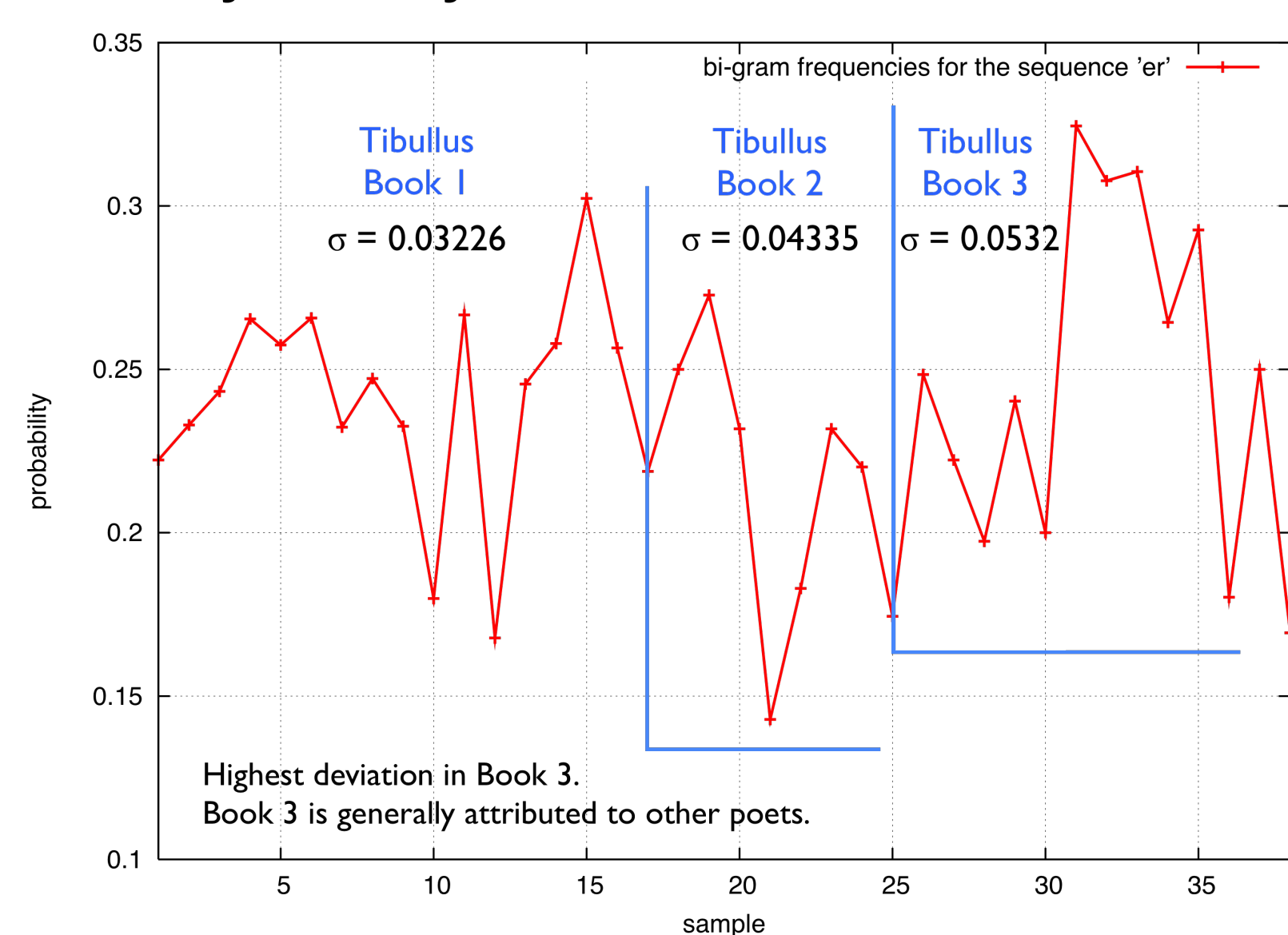
Analysis of stylistic difference in the elegiac couplets of Catullus



The values taken on by a distinct functional n-gram have been found to vary by meter and poet. They can reveal much about the style of a single poet.

Calculating the associated probabilities for *er* over a collection of 50 line samples spanning the entire Catullan corpus exposes a clear break between the elegiac poems (65 - 116) and the rest.

Analysis of stylistic variation between different books of Tibullus



The standard deviation of the bi-gram frequency *er*, calculated over samples drawn from a particular poet, indicates the additional presence of an author signal.

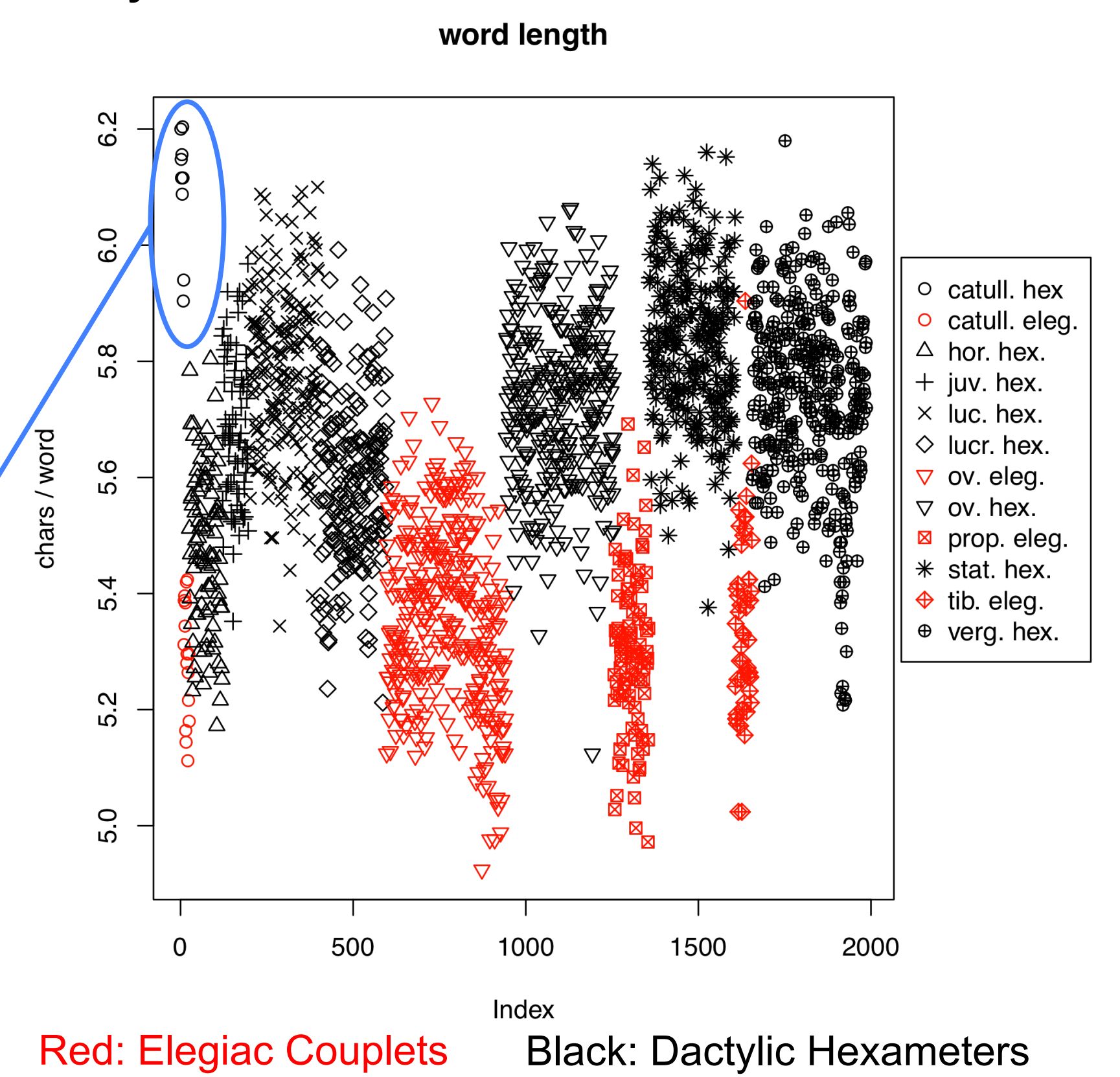
For instance, for 50 line samples representing the three individual books of Tibullus the highest standard deviation belongs to Book 3, which is attributed to a collection of poets, including Tibullus, Sulpicia, and other (often inferior) writers.

A Comparison of Two Meters

Word Length in Elegiac Couplets and Dactylic Hexameters

Beyond bi-gram frequencies, useful results were obtained from mean word length, the feature most sensitive to meter. The number of characters per word tended to be higher for dactylic hexameter than for elegiac couplets both within and between authors.

Catullus 64 was dramatically higher, separated completely from the rest of the Catullan corpus, and generally higher than samples from any author in either meter.



Problem: a deficiency was the lack of a large data base of poets who wrote in both meters.

Solution: split the elegiac corpus into two halves, a hexameter half and a pentameter half, cutting each couplet in two.

A preliminary study considered: Catullus, Ovid, Propertius, and Tibullus.

Samples of 150 randomly-chosen words.

Features considered: the bi-gram frequency *nt*, the ratio *um:am*, and word length

Results: all features are sensitive to the difference between hexameter and pentameter. While, as expected, word length was greater for the hexameter half of the elegiac couplet than for the pentameter, it was still not as high as for stichic (continuous) hexameters. One model to explain this postulates blending of a genre-dependent signal with the meter signal.

References

- C. Forstall, S. Jacobson and W. Scheirer, "Evidence of Intertextuality: Investigating Paul the Deacon's *Angustae Vitae*," presented at Digital Humanities, July 2010.
- N. Coffee, J. Koenig, S. Poornima, C. Forstall and R. Ossewaarde, *The Tesserae Project*. <http://tesserae.caset.buffalo.edu>
- C. Forstall and W. Scheirer, "Features from Frequency: Authorship and Stylistic Analysis Using Repetitive Sound," Chicago Colloquium on Digital Humanities and Computer Science, 2009.
- M. Platnauer, *Latin Elegiac Verse: A Study of the Metrical Usages of Tibullus, Propertius & Ovid*. Cambridge University Press, 1951.
- G. Conte, *Latin Literature: A History*, Translated by J.B. Solodow, the Johns Hopkins University Press, 1999.