A Statistical Study of Latin Elegiac Couplets

C.W. Forstall1 and W.J. Scheirer2
1. Department of Classics, State University of New York at Buffalo
2. Department of Computer Science, University of Colorado at Colorado Springs

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 poetry.

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 of Buffalo’s Tesserae project, an online tool which provides scholars studying Latin poetry easy access to sophisticated textual analysis tools.

This work is part of an ongoing study2 of repetitive sound and its relationship to style in poetry.

Elegiac Couplets

The elegiac meter is used for a variety of themes, most notably Love3. The elegiac couplet is a pair of two different one-line “verses”:

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

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

Latin Elegists considered

Catullus
Ovid
Propertius
Tibullus

Other Latin poets considered

Horace
Juvenal
Lukan
Lucretius
Status
Vergil

Functional n-grams for elegiac couplets:

\( \nu \) – top bi-gram that is common to all poets considered
\( \mu \) – bi-gram with the greatest metrical variation
\( \omega \) – bi-gram sensitive to meter signal

The Significance of the bi-gram er

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.

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.

References