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 literature, including poetry. Existing feature sets and analysis techniques have most often

The Significance of the bi-gram er

In this work, we look at the role repetitive sound plays in the Latin elegiac couplet, where just a

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

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

in this study:
Horace
Juvenal
Lucan
Statius

in this study:

Functional n-grams for elegiac couplets:

e = top bi-gram that is common to all poets considered
b = bi-gram with the greatest metrical variation
s = bi-gram sensitive to meter signal
m = bi-gram sensitive to meter signal

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 Catullian corpus, and generally higher than samples from any author in either meter.

References