Euterpe’s Hidden Song: Patterns in Elegy
Walter J. Scheirer1,2, Christopher W. Forstall3,4
1Harvard University | 2 Univ. of Colorado Colorado Springs | 3Université de Genève | 4Univ. at Buffalo

Aspects of Poetic Composition

“Science is founded on the insight that that world of appearances tells of hidden things, things which appearances reflect but do not resemble. One such insight is that what seems to be disorder in language hides an underlying order—a wonderful order.”

- Gustave Guillaume, *Foundations for a Science of Language*

What is the theoretical model for the patterns we find in language?

**Intentionality**

“Love’s not Time’s fool...”

“But thy eternal summer shall not fade”

“Thine eyes I love...”

**Statistical Learning**

Love

“Love’s not Time’s fool...”

“Thine eyes I love...”

Yule-Simon

“er”

Death

“Death’s second self...”

“On Helen’s cheek all art of beauty set”

In poetry, meter acts as a structuring agent, which tends to overemphasize the underlying natural language statistics that are found in all types of writing and speech. By examining the most commonly occurring elements, a sense of the statistics can be gained with a sampling that overcomes noise.

Case study: we consider the elegies of five major Latin poets and a selection of works in other meters from the Perseus digital library.

Statistical analysis: (1) sound via the functional bi-gram “er” (what the listener hears the most); (2) the interplay between meter and word choice.

**What is the Elegiac form?**

- The first line is identical to dactylic hexameter
- The second, “pentameter,” line of the couplet is shorter by two half-feet

```
[ ] [ ] [ ] [ ] [ ] X
[ ] [ ] [ ] [ ] [ ]
```

*“-” represents a long syllable, “.-” a short syllable, “----” two short syllables or one long, and “X” one long or one short. There is an obligatory word break at “T”.

**Most Frequent Sound Analysis**

Given just the bi-gram “er” for a sample, it is not difficult to guess its meter:

<table>
<thead>
<tr>
<th>Probability</th>
<th>0.45</th>
<th>0.4</th>
<th>0.35</th>
<th>0.3</th>
<th>0.25</th>
<th>0.2</th>
<th>0.15</th>
<th>0.1</th>
<th>0.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>er</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>ls</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>re</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>es</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>un</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| p < 0.0001 |

- Catullus 1-60
- Catullus 61-64
- Catullus 65-116

(a) Books 3 & 4 of Tibullus are generally acknowledged to be the work of others.

(b) Pentameter line ends with a two-syllable word in books 3 & 4 of Propertius.

**Meter’s Influence on Word Choice**

<table>
<thead>
<tr>
<th>Words/Line</th>
<th>0.5</th>
<th>6.0</th>
<th>6.5</th>
<th>7.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characters/Line</td>
<td>30</td>
<td>35</td>
<td>38</td>
<td>40</td>
</tr>
</tbody>
</table>

- elegiac "pentameter"
- elegiac "hexameter"
- dactylic hexameter

The meter’s role in composition:

- Analysis of 100,000 hexameter and elegiac verses from thirteen Latin poets
- Elegiac hexameters have similar numbers of characters to regular hexameter lines, but they’re divided among more words that are shorter
- Blending of genre-dependent signal with meter signal in elegiacs
- The pentameter line tends to have shorter words due to meter; this influences word choice towards shorter words in the hexameter line.

**Software, Data & References**

Feature Generation Code, Source Texts and Spreadsheets:
http://github.com/lesseras/