

**Christopher W. Forstall**  
State University of New York at Buffalo

**Walter J. Scheirer**  
University of Colorado at Colorado Springs

## Introduction

In this work we are looking for new ways to identify stylistic heterogeneity within the Iliad and Odyssey. As oral-formulaic poetry, the Greek epics may contain special evidence of the mutual relationships between poetics, cognition, and creativity.<sup>1</sup> At the same time, scholars of the digital humanities have long recognized that successful digital criticism will find ways to return from statistics to more subjective understanding.<sup>2</sup>

Here, we assign n-gram counts to red, green, and blue color components in order to visualize patterns of sound within the poems.<sup>3</sup> The resulting images demonstrate viscerally that several well-known “set-piece” episodes within Homer’s epics have distinct n-gram distributions.

### Text, sampling, and controls

- The Iliad and Odyssey were downloaded from the Perseus Project<sup>4</sup> in XML.
- They were concatenated, then broken into 20-line samples:
  - once without alteration, (**original** series);
  - 10 more times, each time randomly re-ordering the lines of the poems before sampling (series **r0...r9**)

## n-gram distribution

To detect which might be the most interesting n-grams, we calculated  $s$ , the number of samples in which a given n-gram occurs.

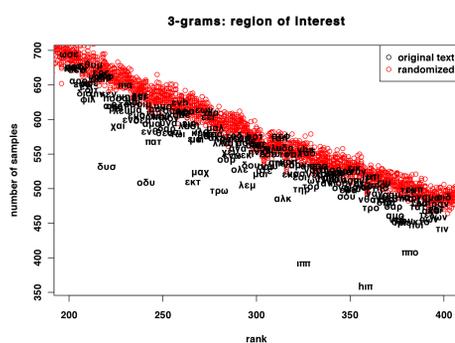
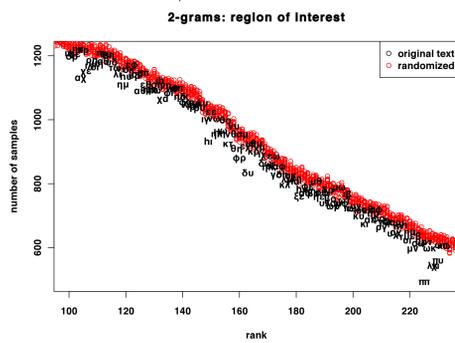
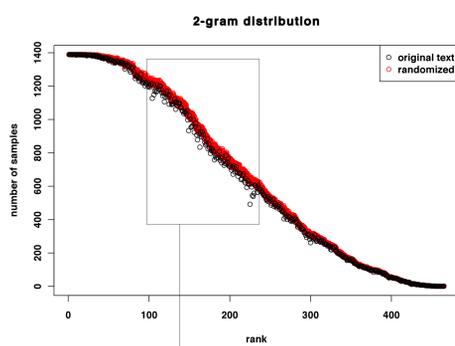
When n-grams cluster in certain samples, other samples go without;  $s$  is then lower in the **original** than in the **r** series. This is seen in the graphs at right.

The lower  $s$ , the more interesting the n-gram.

We quantify this by

$$\text{interest} = s_{\text{original}} - \text{mean}(s_{r_0} \dots s_{r_9})$$

Exactly how interesting this is depends on the variability of  $s$ , so we also consider the standard deviation of  $(s_{r_0} \dots s_{r_9})$ .



top 2-grams, by interest

n-gram	interest	sdev( $s_i$ )
πππ	-146	8
δου	-105	9
αχ	-101	9
ηι	-100	10
χι	-90	7
λκ	-86	7
φρ	-75	12
μν	-70	10
κτ	-69	15
ττυ	-68	5

top 3-grams, by interest

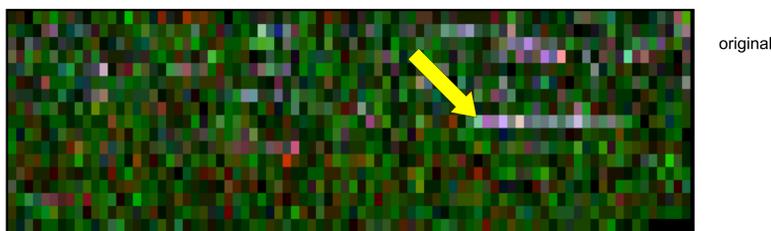
n-gram	interest	sdev( $s_i$ )
ιπππ	-163	7
ηιππ	-162	8
δουσ	-139	6
οδου	-135	11
εκτ	-109	11
τρω	-101	11
ππτο	-100	9
αχι	-98	5
μσχ	-94	5
χιλ	-94	4

## Example I: Three 2-grams for “horse”

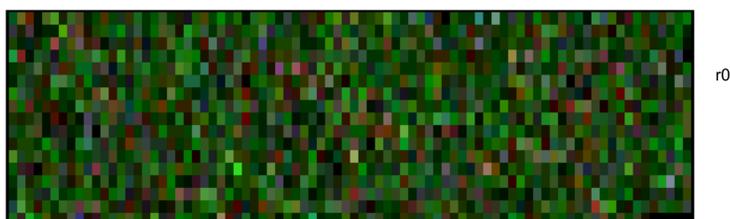
Two of the most interesting 2-grams, **ηι** and **πππ**, are part of the word ἵππος, “horse.” To these we added **ιππ**, (interest = -10; stdev ( $s_i$ ) = 3).

n-gram counts were scaled and translated into color values:

red **ηι**  
green **πππ**  
blue **ιππ**



The large bright region in series **original** corresponds to a set-piece, the chariot race held during the funeral games for Patroclus beginning at Iliad 23.259.



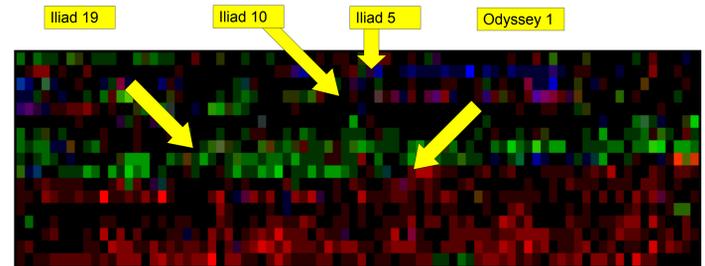
## Notes

1. Peabody, B. (1975) *The Winged Word: A Study in the Technique of Ancient Greek Oral Composition as Seen Principally Through Hesiod's "Works and Days."* Albany: SUNY Press.  
2. Noted already by Packard, D. W. (1947) “Sound Patterns in Homer,” *Transactions of the American Philological Association* 104:239–260.

## Example II: Three heroes

Next we consider three 3-grams related to three independent content elements. Each is a component of a Greek hero’s name:

red **δουσ** Odysseus  
green **χιλ** Achilles  
blue **τυδ** Diomedes<sup>6</sup>

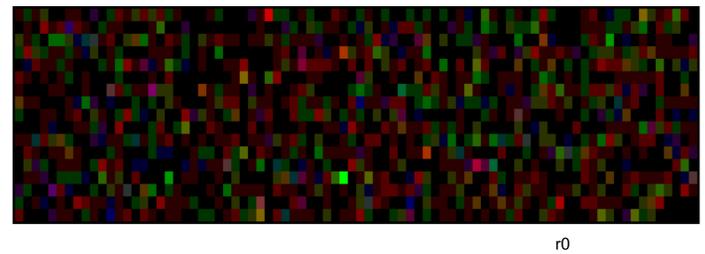


Each hero is foregrounded in a different part of the story:

- Achilles in Iliad 19 and following
- Odysseus in the Odyssey
- Diomedes in Iliad 5

Note the purple section in Iliad 10. This is the night raid in which Diomedes and Odysseus work together.

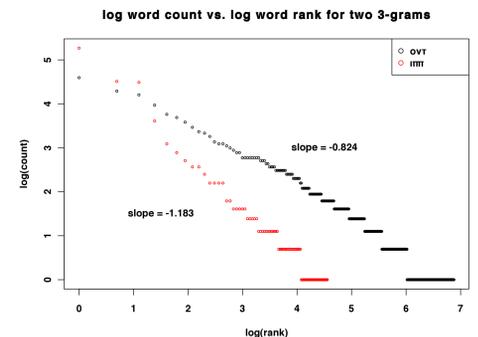
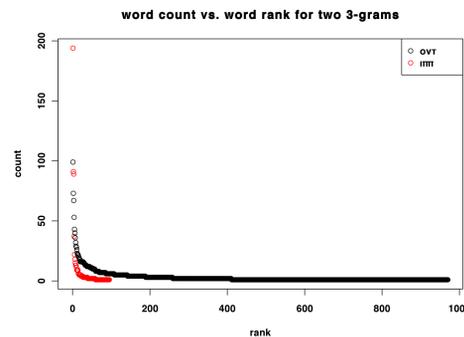
Not only do the colors show independent thematic elements; they also can represent their interactions.



## Quantifying the sound-content relationship

We measure both the number of words containing the n-gram, and the number of times each of those words occurs. The greater the lexical diversity of an n-gram, the less content-driven it is likely to be.

For example, compare the 3-grams **ιπππ** and **οντ**, both frequent and of high interest values.



The graphs above show the number of times each word containing a given n-gram occurs in the text as a function of that word’s rank. The right is a log-log version of the left. The **οντ** curve has a shallower slope.

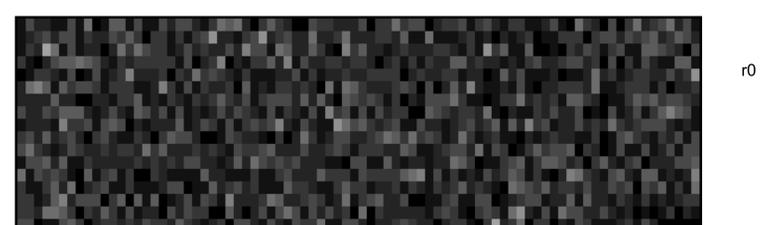
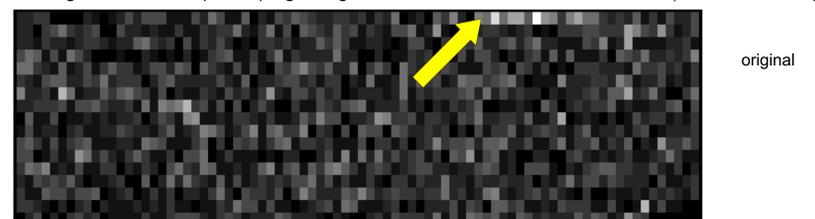
Below are the top 10 words for each. **οντ** shows a far greater diversity than **ιπππ**, in large part because it contributes to some common noun and verb inflections.

top words containing <b>ιπππ</b>			
rank	count	word	meaning
1	194	ἵππους	“horses” (acc. pl.)
2	91	ἵππων	“horses” (gen. pl.)
3	89	ἵπποι	“horses” (nom. pl.)
4	37	ἵππότης	“horseman”
5	22	ἵπποδάμοιο	“horse-tamer” (gen. s.)
6	18	ἵπποισιν	“horses” (dat. pl.)
7	15	ἵππον	“horse” (acc. s.)
8	13	ἵπποισι	“horses” (dat. pl.)
9	13	ἵππων	“horses” (nom./acc. d.)
10	11	ἵππηλάτα	“chariot driver”

top words containing <b>οντ</b>			
rank	count	word	meaning
1	99	ἔοντα	“being”
2	73	πόντον	“sea” (acc. s.)
3	67	ἔποντο	“they followed”
4	53	ἔχοντες	“having” (nom. pl.)
5	43	πόντωι	“sea” (dat. s.)
6	40	ἴκοντο	“they arrived”
7	36	ἔγένοντο	“they were”
8	32	ἔχοντα	“having” (acc. s., nom./acc. pl.)
9	29	ἔοντες	“going”
10	28	γέροντος	“elder”

## Example III: A content-independent pattern?

Here, we used shades of grey to represent counts of a single 3-gram, **οντ**. The bright region corresponds to the “Catalogue of Ships,” Iliad 2.484–759, and provides a subjective appreciation of the sound patterns that distinguish this inset piece (beginning with its own invocation to the Muses) from the main poem.



3. We take our inspiration in part from Plamondon, M. (2009) “Computational Phonostylitics: Computing the Sounds of Poetry,” presented at DHCS 2009.  
4. Perseus Digital Library Project. Ed. Gregory Crane. <http://www.perseus.tufts.edu>. Accessed 2/1/2010.  
5. In composing the n-grams, we transcribe rough breathing as Latin **h**; iota subscript as adscript (regular **i**); and final sigma with medial sigma as **σ**.  
6. By way of his patronym, “Son of Tydeus.”