

What makes an Allusion? A Digital Approach

Christopher W. Forstall,¹ Neil Coffee,¹ James Gawley,¹ and Walter Scheirer²

¹ The University at Buffalo, State University of New York; ² University of Colorado, Colorado Springs

<http://tesserae.caset.buffalo.edu>

What is an allusion?

This is a long-standing question in literary and linguistic studies. By training a computer system to detect the allusions that readers find between two texts, we arrive at formal, measurable characteristics of allusion—and discover new intertexts that traditional researchers have overlooked.

Tesserae detects the re-use of distinctive two-word phrases across texts, accounting for spelling differences and inflection.

English: William Wordsworth's *Prelude*, Book 1 vs. William Cowper's *The Task*

	target phrase	source phrase	matched on	score
1.	wordsworth prelude 1.233 Thoughtfully fitted to the Orphean lyre ;	cowper task 3.587 By magic summons of the Orphean lyre .	lyre, orphean	9
2.	wordsworth prelude 1.88 A backward glance upon the curling cloud	cowper task 4.472 There sit involved and lost in curling clouds	cloud, curl	9

Greek: Apoloniou's *Argonautica* vs. Homer's *Illiad*

	target phrase	source phrase	matched on	score
1.	A.R. 1.850 ῥηιδίως. Κύπρις γὰρ ἐπὶ γλυκὺν ἴμερον ὤρσεν	hom. il. 3.139 ὣς εἰποῦσα θεὰ γλυκὺν ἴμερον ἔμβαλε θυμῷ	ἴμερος, γλυκὺν	9
2.	A.R. 1.927 στησάμενοι κούρης Ἀθαμαντίδος αἰπὰ ῥέεθρα	hom. il. 8.369 οὐκ ἂν ὑπεξέφυγε Στυγὸς ὕδατος αἰπὰ ῥέεθρα .	ῥεῖθρον, αἰπὰ	9

Latin: Vergil's *Georgics* vs. Lucan's *Bellum Civile*

	target phrase	source phrase	matched on	score
1.	verg. g. 1.495 exesa inveniet scabra robigine pila	luc. 1.243 Et scabros nigrae morsu rubiginis enses.	robigo, scaber	9
2.	verg. g. 1.315 frumenta in viridi stipula lactentia turgent.	luc. 3.503 Nec, quamvis viridi lactetur robore, lentas	lacteo, uiride-uiridis	9

The Tesserae working-definition of allusion

Two-word (bigram) matching by lemma gives first-stage results. These are then ranked by the proximity of the words in the two-word phrases and the rarity of those words.

$$\text{score} = \ln \left(\frac{\sum \frac{1}{f(t_i)} + \sum \frac{1}{f(s_i)}}{d_t + d_s} \right)$$

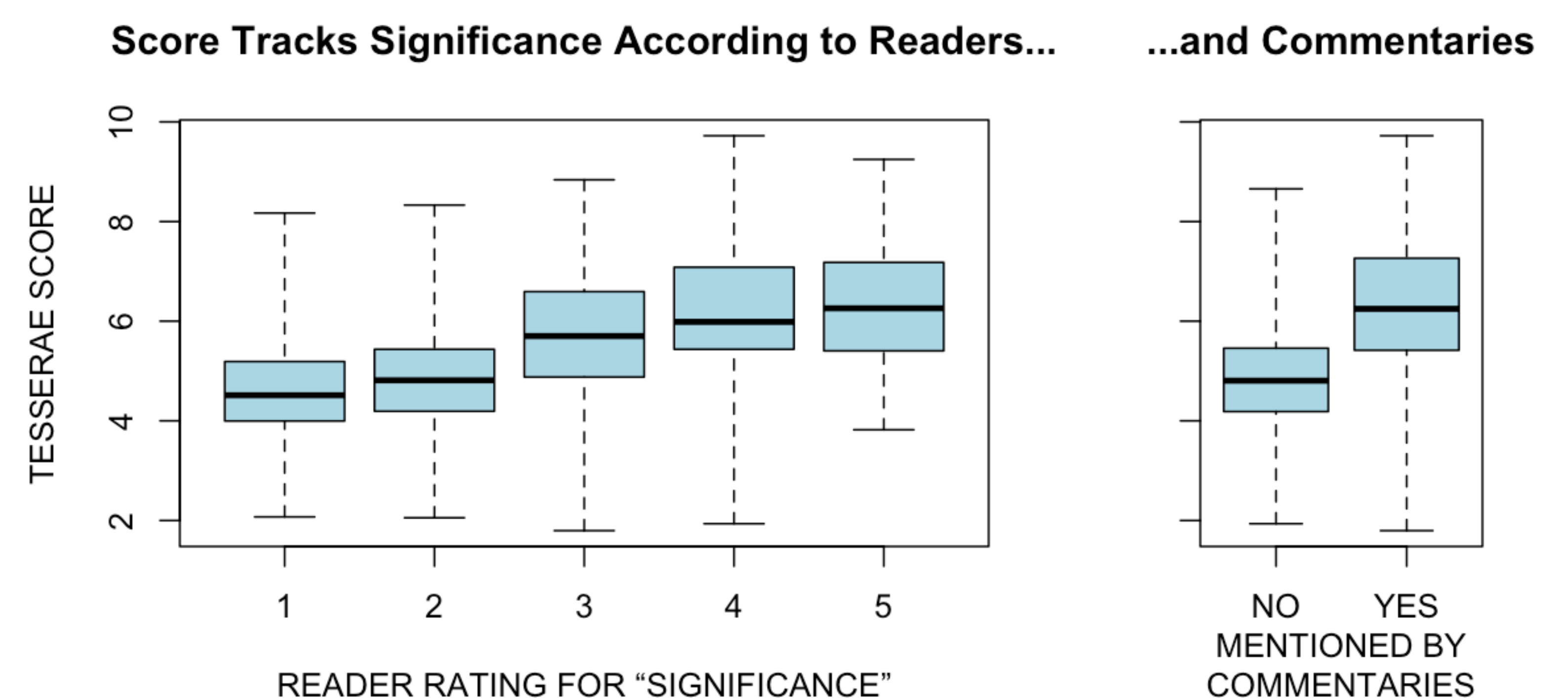
Where,

$f(t)$ is the frequency of each matching term in the target phrase;
 $f(s)$ is the frequency of each matching term in the source phrase;
 d_t is the distance in the target;
 d_s is the distance in the source.

“Distinctiveness” in allusion is a function of the *frequency* and *proximity* of the two words in the matching phrase. When rare words appear close together, we are confident of an allusion.

How do Tesserae results compare to traditional scholarship?

Testing on a sample of Latin epic poetry shows that Tesserae identifies a substantial majority of the allusions important to traditional scholarship.



The figures above represent approximately 3,000 instances of parallel phrases drawn from Lucan Civil War Book 1 and Vergil Aeneid, meeting the minimum criterion for allusion, i.e. two or more words shared between Lucan and Vergil. The reader ratings were produced by graduate student and faculty readers during a seminar on intertextuality at the University at Buffalo. Commentary data was drawn from Heitland and Haskins (1887), Thompson and Bruère (1968), Viansino (1995), and Roche (2009).

What about more complex allusions?

Not all allusions are characterized by the re-use of two-word phrases. In order to detect more subtle intertextuality, modules are currently under development to match phrases based on **synonyms**, **sound patterns**, and **themes**. At first these characteristics will be used to help determine the distinctiveness of a two-word match; at a later stage it may be possible to detect matches on these criteria alone.



TESSERAЕ



University at Buffalo
The State University of New York



University of Colorado
Colorado Springs



NATIONAL ENDOWMENT FOR THE
Humanities