

CESTA research anthology 2022

This is an excerpt from the *CESTA Research Anthology 2022*. To read the full *Anthology*, [click here](#). To learn more about CESTA's *Anthology* series and read editions from previous years, [click here](#).

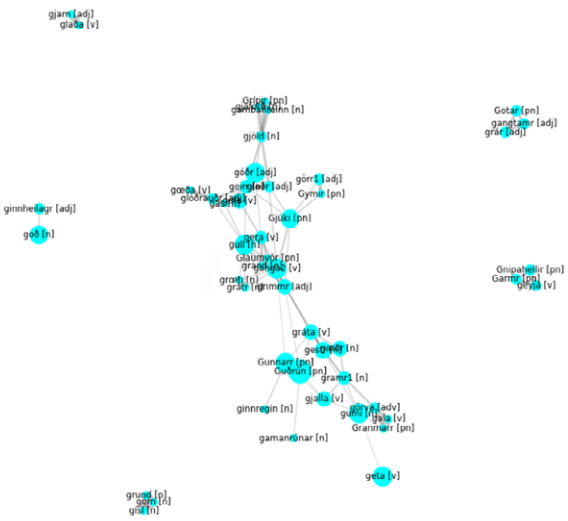
Raiding the Wordhoard: Recurring Alliterative Collocations in Old Norse Eddic Poetry

by James Parkhouse (Visiting Scholar, Text Technologies)

My project examined recurring alliterative collocations in Old Norse eddic poetry, a corpus of around 12,000 verses of mostly anonymous mythological and legendary poetry composed in medieval Scandinavia and associated colonies. Old Norse poetic meter was predicated on a pattern of alliterating stressed syllables per line, rather than on a fixed number of syllables. This gave rise to the repeated combination of alliterating words; as illustrated by examples in modern English such as 'hearth and home' or 'life and limb', such expressions can have rhetorical resonances beyond their semantic significance; the study of collocations in Old Norse poetry can thus offer valuable insights into the aesthetics of the texts and the worldviews of their composers and audience. However, since there are finite numbers of words beginning with any given letter, some combinations could recur by coincidence.

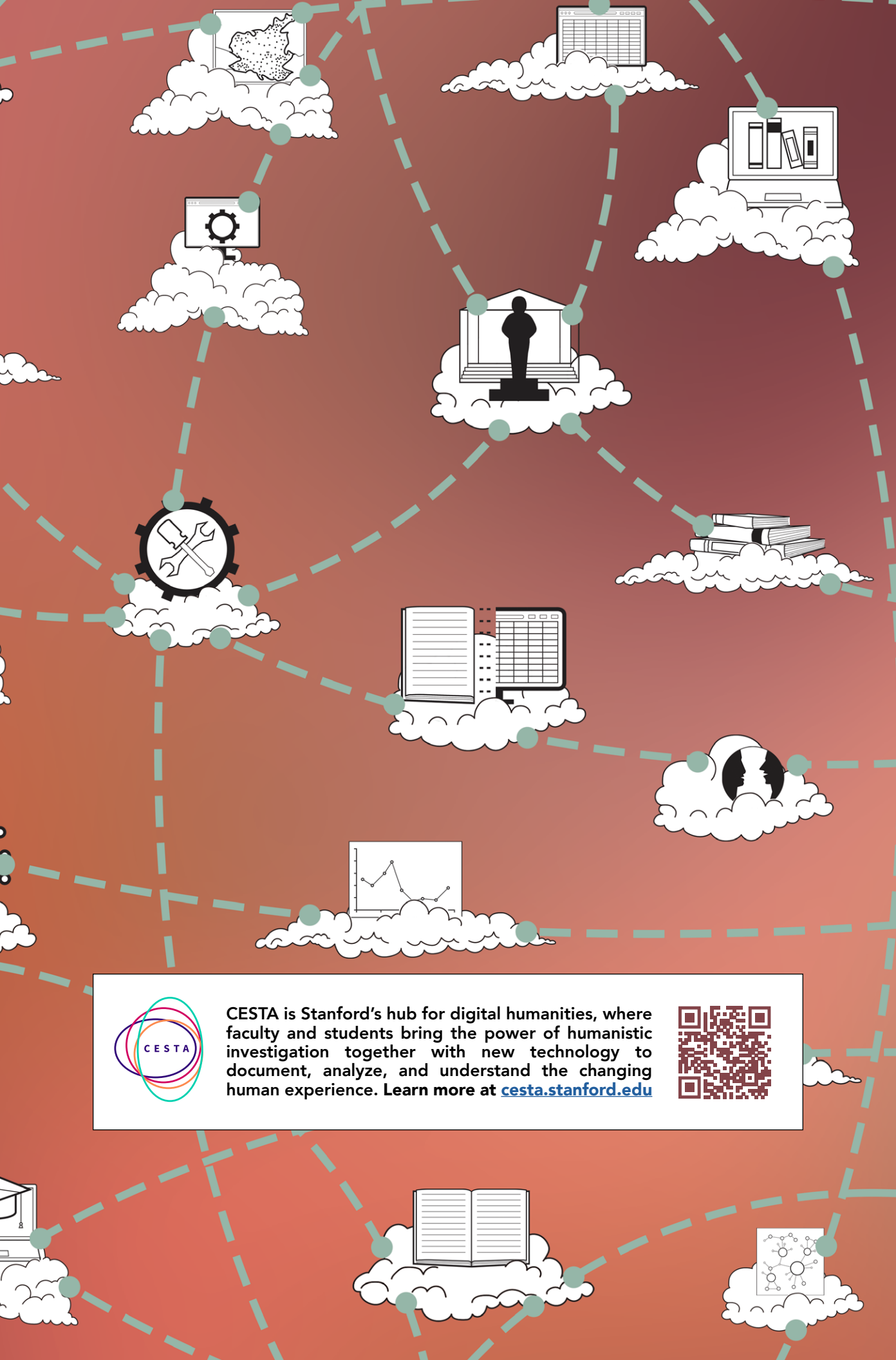
With the assistance of my intern, Poojit Hegde, I employed statistical testing to identify which alliterative combinations recurred more often in the corpus than would be expected by chance, indicating that these were deliberate poetic devices. (Due to time constraints, I conducted a proof-of-concept on one important manuscript, comprising about half of the total corpus). I used Fisher’s exact test, which evaluates the independence of two variables given their observed distribution. I first compiled a spreadsheet of the alliterating stresses in each verse of the corpus. I cleaned this raw data of corrupt verses (those with defective meter or indecipherable meaning as a result of centuries of manuscript transmission), and lemmatized inflected forms to the root form. I then used Python to generate contingency tables for each collocation with a frequency of two or more, and to perform mass statistical testing.

A conservative correction to avoid false positive results gives a significance threshold of $p = 0.00006$; 77 collocations met this threshold, providing a springboard for literary-critical analyses of the poetic functions of these combinations. Additionally, I have begun experimenting with visualizations of wider collocational networks using the open-source graphic software Gephi, treating each collocote as the focal node of an ego network.



An example of work in progress visualizing collocational networks for the initial letter g-.





CESTA is Stanford's hub for digital humanities, where faculty and students bring the power of humanistic investigation together with new technology to document, analyze, and understand the changing human experience. Learn more at cesta.stanford.edu

