Linking Epic Speeches

Forstall, Christopher W.
cforstall@mta.ca
Mount Allison University, Canada

Stagg, Wyatt
wastagg@mta.ca
Mount Allison University, Canada

In this poster we present a suite of Python-based tools for interacting with DICES, an open dataset for the diachronic study of direct speech in Greek and Latin epic poetry (Forstall et al. 2022). The *dices-client* package not only permits programmatic search of the DICES database but more importantly supports its integration within an ecosystem of linked open data (LOD).

The narrative tradition of Greco-Roman epic poetry is expansive and deeply intertextual (see, e.g., Hinds 1998 and Hardie 1993). Homeric characters from the eighth century BCE appeared in "sequels" by Virgil in the first century BCE and by Quintus of Smyrna in the fifth century CE. The Argonauts' quest for the golden fleece was reinvented by Apollonius in Hellenistic Egypt and again by Valerius Flaccus in Flavian Rome. To study the evolution of this genre thus demands a similarly collaborative approach. The DICES database and the client package presented here were designed to leverage existing LOD resources for digital classics, facilitating analysis of the use of character speech across periods and languages within a single scripting environment.

We demonstrate the search functionality of the DICES python client in comparison with a more traditional web-based user interface, highlighting support for distant reading approaches and presenting summary statistics of the underlying DICES dataset. Then, in a constellation of short case-studies, we explore how the client's supporting modules provide direct interaction between speech records and several classes of external digital tools:

- Speaker and addressee identities are cross-referenced against two databases of named entities: WikiData, a publicly-editable general-purpose knowledge base, and MANTO, a curated scholarly collection of mythological entities from Greek and Latin literature (Hawes / Smith 2022).
- Records on textual passages can be linked to digital libraries using the Canonical Text Services protocol (Tiepmar / Heyer 2019) in order to retrieve the full text of the speech or additional metadata about the passage, work, or author.
- Retrieved passages can further be processed using the natural language processing pipelines provided by the Classical Language Toolkit (CLTK), including tokenization, stemming, part-of-speech-tagging, and syntactic parsing (Burns 2019).

For example, we can query DICES to see how often transtextual characters such as Achilles speak in epics from Homer to Quintus (Table 1). Then, from the same script we can also request the text of Achilles' speeches from the Perseus Digital Library and pass them directly to CLTK to see how the swift-footed hero uses different parts of speech across authors and languages (Figures 1, 2).

Drawing on the character metadata in MANTO and WikiData, we can explore direct speech through the lens of relationships or other personal attributes not necessarily present in the speeches themselves: for example, we can compare the language used between mothers and sons across the corpus, or the conversational patterns of characters destined to die at the hands of their interlocutors.

Working examples of these case studies as well as additional tutorial materials are available as Jupyter Notebooks on GitHub: https://github.com/cwf2/dices-examples. The DICES client itself is available at https://github.com/cwf2/dices-client.

Table 1. Distribution of speeches by selected transtextual characters

	Achilles	Aeneas	Diomedes	Helen	Heracles	Nestor	Odysseus
Apollonius, *Argonautica*	0	0	0	0	2	0	0
Homer, *Iliad*	88	6	26	7	0	32	27
Homer, *Odyssey*	3	0	0	4	1	9	169
Lucan, *Civil War*	0	0	0	0	1	0	0
Nonnus, *Dionysiaca*	0	0	0	0	0	0	0
Ovid, *Metamorphoses*	5	1	1	0	8	4	3
Quintus, *Posthomerica*	8	1	5	2	0	14	8
Silius Italicus, *Punica*	0	1	1	0	1	0	0
Statius, *Achilleid*	7	0	3	0	0	0	7
Statius, *Thebaid*	0	0	0	0	1	0	0
Theocritus, *Idylls*	0	0	0	0	3	0	0
Triphiodorus, *Sack of Troy*	0	0	0	0	0	0	1
Valerius Flaccus, *Argonautica*	0	0	0	0	6	1	0
Virgil, *Aeneid*	0	70	1	0	0	0	0

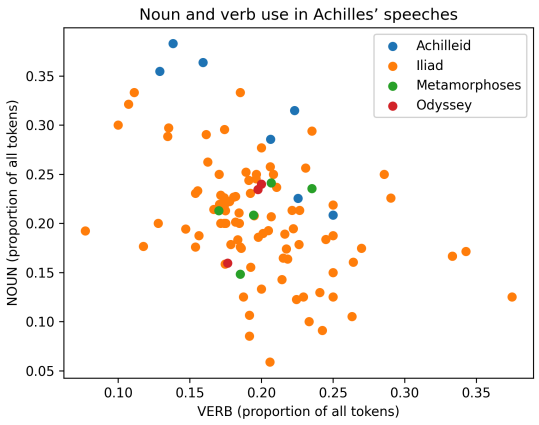

Figure 1. Achilles' use of nouns vs. verbs in his speeches, in Homer's Iliad and Odyssey, Ovid's Metamorphoses, and Statius' Achilleid

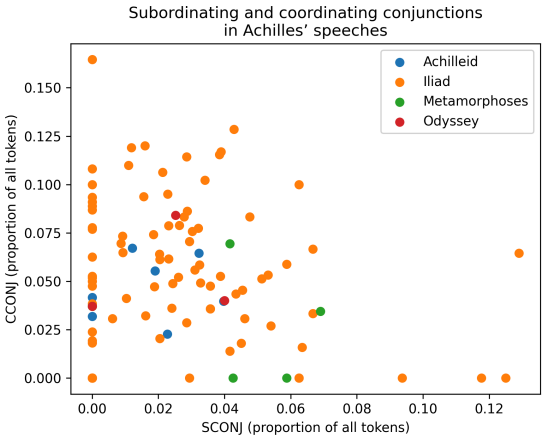

Figure 2. Achilles' use of subordinating vs. coordinating conjunctions in Homer, Ovid, and Statius. We had initially expected the largest difference to be between languages, but in fact, the more significant separation was between the two Roman authors.

Bibliography

Burns, Patrick J. (2019): "Building a text analysis pipeline for classical languages", in: Berti, Monica (ed.) *Digital Classical Philology: Ancient Greek and Latin in the Digital Revolution*. Berlin: De Gruyter 159–176.

Forstall, Christopher W. / **Finkmann, Simone** / **Verhelst, Berenice** (2022): "Towards a linked open data resource for direct speech acts in Greek and Latin epic", in: *DSH* 37, 4: 972–981.

Hardie, Philip (1993): *The Epic Successors of Virgil: A Study in the Dynamics of a Tradition*. Cambridge: Cambridge University Press.

Hawes, Greta / **Smith, Scott** (2022): "A dataset of mythical people with stable URIs." https://www.manto-myth.org/blog/a-dataset-of-mythical-names-with-stable-uris [2022.11.04]

Hinds, Stephen (1998): *Allusion and Intertext: Dynamics of Appropriation in Roman Poetry* . Cambridge: Cambridge University Press.

Tiepmar, Jochen / **Heyer, Gerhard** (2019): "The Canonical Text Services in Classics and beyond", in: Berti, Monica (ed.) *Digital Classical Philology: Ancient Greek and Latin in the Digital Revolution*. Berlin: De Gruyter 95–114.