Corpus-linguistic approaches to phonological wordhood in metrical corpora

Dieter Gunkel (University of Richmond) & Kevin M. Ryan (Harvard University)
dgunke@richmond.edu & kevinryan@fas.harvard.edu

The close study of corpora composed in poetic meter can yield a wealth of information about the prosodic phonology of a language, as demonstrated in Devine and Stephens’ *The Prosody of Greek Speech* (1994), to mention an outstanding example. In this study, we develop several corpus-linguistic methods for studying clisis and phonological wordhood, which we apply to the Ṛgveda, a 164,766-word corpus of archaic Sanskrit hymns composed in poetic meters that regulate syllable count and syllable weight as well as the distribution of phonological word and phrase boundaries.

One new method estimates the clisis of a particular morphological word, e.g. *ca* ‘and’, by comparing the distribution in verse of *ca* and a preceding morphological word, e.g. *indraś ca* ‘and Indra:NOM.SG’, with the distribution of single morphological words of the same phonological shape, e.g. *indrasya* ‘Indra:GEN.SG’. Another method uses adjacency to metrical boundaries (caesura, verse boundary, etc.) to estimate clisis.

Combining these methods, we locate morphological words along a cline of clisis (thus estimated), then investigate the phonological and morphosyntactic factors that contribute to it. Our approach is meant to complement (not to supersede) existing approaches to the study of Ṛgvedic clisis and phonological wordhood, e.g. the study of sandhi (e.g. Oldenberg 1888; Hale 1990, 1995) and word order (e.g. Hale 2007, Lowe 2014).

References


