Computational Phonostylistics: Computing the Sounds of Poetry

Ian Lancashire suggests that the fields of cognitive psychology and neuroscience "are enhancing our understanding of how authors create both oral and written texts." Recent work in neuroscience suggests that phonemes, or, at least, the muscle impulses required for the production of phonemes, are at the basis of our brain's ability to produce communication. Dr. Philip Kennedy is working on giving synthesized speech to patients suffering from "locked-in syndrom" by picking up the brain's signals that otherwise would produce phonemes and having a computer program sort through the signals and translate them into sound (Baker 52, 56). This research suggests that communication, including writing, is phoneme-based and lends strong support to what many have been suspicious of for centuries: that literary production is not purely semantic, that the phonological content–the sounds–of texts must have significant importance.

My research is in the field of computational phonostylistics, the identification of patterns and other characteristics in the phonological content of texts. I focus in particular on poetry, with the ultimate goal of characterizing different poetic styles–of author and of time periods–and perhaps also quantifying poetic beauty. Poetry possesses a significant characteristic that prose usually lacks: it exists in two dimensions. The first dimension is of course the sequential reading of the words, or the sounds of the words, of the poem (the left-to-right dimension); the second dimension is the arrangement of the words on the page, which of course incorporates the first dimension but adds the second (the top-to-bottom dimension).

My presentation will outline some of the theoretical models and computational techniques I have adopted for my research into the phonostylistics of English poetry. In particular, it will present a technique for cluster analysis of poems based on a two-dimensional representation of poetry, where the phonemes of a line influence not just the adjacent phonemes, but also the phonemes above and below them in the poetic lines. It will also present my theory of phonemic persistence and the calculation of phonemic accumulations. The theory argues that a phoneme's effect on the reader will carry through to the subsequent phonemes (on a one-dimensional level), and the phonemic accumulations are an attempt to quantify the effect phonemes (or phoneme classes) have upon the reader. The resulting data can be treated as a phonemic accumulation waveform.

The presentation will also outline some of the numerical analysis techniques that have already been adopted and those that have yet to be adopted. The phonemic accumulation waveforms can be processed as phonemic signals, where the "noise" can be identified and filtered, resulting in what is theoretically an idealized representation of the phonemic effects of the reading experience. The addition of two or more signals, resulting from two or more different classes of phonemes, can yield interference graphs, where the competing effects of classes of phonemes can be modelled for any given syllable in a poem. The potential for advances in textual stylistics will be outlined.

Works Cited

Baker, Sherry. "Rise of the Cyborgs." *Discover*. October 2008. 50-57.

Lancashire, Ian. "Cognitive Stylistics and the Literary Imagination." *A Companion to Digital Humanities*. Eds. Susan Schreibman, Ray Siemens, John Unsworth. Oxford: Blackwell, 2004. <http://www.digitalhumanities.org/companion/>