Chapter 16

Computer-Aided Rhetorical Analysis

Suguru Ishizaki
Carnegie Mellon University, USA

David Kaufer
Carnegie Mellon University, USA

ABSTRACT

This chapter presents a corpus-based text analysis tool along with a research approach to conducting a rhetorical analysis of individual text as well as text collections. The motivation for our computational approach, the system development, evaluation, and research and educational applications are discussed. The tool, called DocuScope, supports both quantitative and quantitatively-informed qualitative analyses of rhetorical strategies found in a broad range of textual artifacts, using a standard home-grown dictionary consisting of more than 40 million unique patterns of English that are classified into over 100 rhetorical functions. DocuScope also provides an authoring environment allowing investigators to build their own customized dictionaries according to their own language theories. Research published with both the standard and customized dictionaries is discussed, as well as tradeoffs, limitations, and directions for the future.

INTRODUCTION

Our research seeks to account for the wide variation in the experiences texts afford. In particular, we are interested in uncovering how the writer’s small and recurring linguistic choices contribute to the whole text experience of the reader. Answering this question is important to rhetorical analysts who wish to understand how the plasticity of language choice affects the plasticity of responses to rhetorical situations. Answers to these questions are also important to writing instructors who wish to understand the wide palette to which students must be exposed in order to master writing across a range of genres and situations. Our research is concerned with the kernel of rhetorical theory by

DOI: 10.4018/978-1-60960-741-8.ch016
exploring the relationship between micro (surface) linguistic choices versus a holistic rhetorical effect in language design. We have sought to explore the extent to which local language decisions traditionally associated with “style” aggregate to inform global linguistic organization associated with “invention” and “genre” (Bawarshi, 2003).

In what sense can micro-selections of textual expressions contribute compositionally to a text’s overall genre features? Research on first and second-language learning has independently converged with and benefited our efforts, particularly the cognitive work of Ellis (Ellis & Ferreira-Junior, 2009), and the linguistic work of Hoey (2005). From different disciplines, both researchers have demonstrated that language learners do not learn a single word at a time, but they acquire strings of words as units over time through a process which Ellis refers to as a “statistical ensemble.” A conventional dictionary enumerates the various meanings and parts of speech (e.g., noun, verb, adjective) a word may reference, but language users learn words by understanding the various streams of discourse to which individual words contribute in a culturally and socially meaningful context. Take the single word “smear.” Language users learn that the expression to “smear a politician” contributes to a negative expression while “smearing soap in the shower” contributes to an expression of everyday motion. Hundreds of millions of everyday strings like these enter a cultural repository of meaning-making potential. Speakers and writers exploit the cultural repository to be understood. Listeners and readers exploit it to understand. Through rhetorical strategies, speakers and writers may invent new expressions, which become candidates for entering and extending the cultural repository. The picture looks something like Figure 1 with respect to writers, readers, and texts.

These language patterns with “smear” and others alone have small but real effects. Taken as an aggregate, they exert great influence on the experience of the reader. At the whole text level, language patterns can determine a number of holistic effects about the overall shape of a text. In other words, what we call types or genres of writing are created by the small and often implicit design decisions writers make along these and many other choices. Our research has sought

---

**Figure 1.** A picture of literate language learning. Writers rely on previous shared experience with language patterns to be understood. Readers rely on this same repository to understand. Writers, through rhetorical strategies that may be site specific, can invent new expressions that may or may not add to the shared cultural inventory.