

A Contrastive Corpus Analysis of Modern Art Criticism and Photography Criticism

Arthur Hullender and Philip M. McCarthy

University of Memphis
304 Patterson
Memphis, TN
{ahullndr, pmmccrth} @memphis.edu

Abstract

In this study, we analyze two corpora of art critiques: one on the subject of photography and the other on the subject of modern art. We use two computational tools, the Gramulator and GPAT to analyze both sets of texts. The Gramulator was used to show the indicative linguistic features that make photography criticism a distinct genre from modern art criticism. Results suggest that lexical features, structural formats, and genre consistency differed significantly between the two corpora. The findings provide information for teachers, students, publishers, and curriculum developers for creating more effective writing and teaching materials. This includes material for English for Specific Purposes (ESP) in the form of textbooks, workbooks and other external learning material.

Introduction

Our study focuses on the language features of *art criticism*. More specifically, we are interested in the differences between language used for *photography criticism*, and the language used for *modern art criticism*. Our research question is “Can the language of photography criticism in terms of indicative linguistic features be considered as a genre distinct from modern art criticism?” And if so, which features of language are driving these differences. To address our research questions, we formed the hypothesis that the language of photography criticism will contain “process specific features” when analyzing photography because, unlike modern art, photography is dependent upon a technical chemical process to produce a composition (Diamond and Weiss 2002). Further, Soykan (2009) argues that the vocabulary used to describe art varies according to the art under analysis so much that the genres can be described as art “languages.”

The goal of this study is to discover and assess the language differences used in photography criticism and modern art criticism, and, based on our findings, to offer some ideas as to the effect these language features might have on the communicative goals of writers, as well as the teaching implications for English for Specific Purposes (ESP). The study is of interest to writers, textbook designers, curriculum designers, and ESL teachers; especially those working in ESP related areas. The study is also important to linguists and cognitive scientists because it stands to explain how differences in perceived categories (Modern Art, Photography) are created through linguistic features.

Corpus

We constructed two contrasting corpora. Both corpora comprised critiques taken from magazines and newspapers related to their interests. The final corpora comprised random text samplings from 94 modern art critiques (MAC) and 48 photography critiques (PC), with each text size approximately 1000 words.

The Tools: The Gramulator and GPAT

The major tool in our study is the computational contrastive analysis software, the *Gramulator*. This software allows the user to identify lexical features that are *indicative* of specific texts (McCarthy, Watanabe, and Lambkin in press). The Gramulator processes both corpora (here PC and MAC) relative to each other, outputting sets of *differentials*, which are features typical of one corpus, but untypical of the other corpus. The differentials in this study take the form of *bigrams*. When treated as an array of features, the differentials form indices. Thus, PC (MAC) represents an array of n-gram differentials that are included if, and only if, they are typical of the PC corpus and untypical of the MAC corpus. Similarly, MAC (PC) is

typical of the MAC corpus but untypical of the PC corpus. Other computational tools such as Coh-Metrix (Graesser et al. 2004) produce only independent analysis of a single corpus. Thus, the Gramulator is particularly well suited to *contrastive analysis*, as is the goal of the current study.

The second software we use is the Genre Purity Assessment Tool (GPAT: McCarthy 2010). GPAT analyzes texts for language elements specific to either the science or narrative genre. McCarthy (2010) demonstrated that GPAT's genre accuracy was at least as high as a combination of over 30 Coh-Metrix measures.

Results

We used the Gramulator's *Sorter* module to place two-thirds of the texts randomly from each corpus into *training sets*, and the remaining texts into *test sets*. We used the Gramulator's main module to analyze the training sets and create indicative indices of each corpus [i.e., PC (MAC) and MAC (PC)]. We conducted t-tests to assess the effect of the indices MAC (PC) and PC (MAC) on the data of each corpus. The results validated the indices. For index MAC (PC): MAC: $M = 0.063$, $SD = 0.015$; PC: $M = 0.028$, $SD = 0.016$; $t(1, 66) = 9.427$, $p < 0.001$, $d = 2.29$. A similar result was found for the PC (MAC) index PC: $M = 0.051$, $SD = 0.015$; MAC: $M = 0.03$, $SD = 0.007$; $t(1.66) = 7.746$, $p < 0.001$, $d = 1.882$.

Having validated our data and approach, we examined the differentials as linguistic features using the concordancer module of the Gramulator. Differences between usage are reported using Fisher's Exact Test. We also looked at combined patterns of *n*-gram collocations to show how *n*-grams are often semantically related if not lexically related. We refer to these combinations of elements as flexigrams.

For the PC corpus, 8 of the highest ranked 15 bigrams (in terms of weighted frequency) included self-referencing words like *photography* and *photograph*. Each of the PC bigrams was found to be significant at $p < .001$. In contrast, the MAC bigrams showed terms related to materials and process. These included *to paint* ($p = .011$) and *the canvas* ($p = .007$). The results also suggested a flexi-gram pattern (i.e., semantically related *n*-grams): the most common flexi-gram being *to paint* + [noun] ($p = .002$) and *the canvas* + [preposition] ($p < .001$).

The MAC corpus also contained bigrams where abstract referencing helped to form contrastive constructs: *kind of*, *in common*, *than in*, *but there*, *without the* (see Table 1 for context). Although only one of these bigrams was individually significant (*in common*, $p = .049$), as a flexi-gram the combined bigrams show that MAC contains language that is a form of hedging, abstracting, or setting up a contrast ($p < .038$). We argue that the critique writers, having a problem making concrete descriptions of the nonverbal image, must compensate by using abstract

language as a contrastive description for the purposes of situating their analyses. We further speculate that such language may help to better form an appropriate mental model to claim the writers' interpretation.

GPAT results also demonstrated a significant contrast between Photography and Modern Art texts. MAC showed 50 out of 97 are narrative (*ns*), but PC showed 31 out of 48 are science ($p < .001$). This confirms our hypotheses that PC contains more science-related features.

Table 1. Flexi-grams of MAC contrastive constructs

Aesthetic distance means separation, a <i>kind of</i> transcendence, if you please.
Magritte's best images have more <i>in common</i> with reporting than with fantasy.
Picasso has seldom been more tender <i>than in</i> his first portrait of Marie-Thérèse Walter.
<i>But there</i> was a more general sense of ferment at black mountain.
Magritte's poetry was inconceivable <i>without the</i> banality on, and through, which it worked.

Discussion

This study provided evidence that there are specific linguistic features that are indicative of both photography criticism and modern art criticism. Taken as a whole, the results support the position that the two text types are distinct genres. This study is a small but important step on the path to a greater understanding of genre classification for art critiques in terms of lexical features.

References

- Diamond, A., and Weiss, D. 2002. *Handbook of Imaging Materials*. CRC.
- Graesser, A. C.; McNamara, D. S.; Louwerse, M.; and Cai, Z. 2004. Coh-Metrix: Analysis of Text on Cohesion and Language. *Behavior Research Methods, Instruments, and Computers* 36:193-202.
- McCarthy, P. M. 2010. GPAT Paper: A Genre Purity Assessment Tool. In Guesgen, H. W. and Murray, C. (Eds.), *Proceedings of the 23rd International Florida Artificial Intelligence Research Society Conference*, 241-246. Menlo Park, CA: AAAI Press.
- McCarthy, P. M.; Watanabe S.; and Lamkin, T. A. in press. The Gramulator: A Tool for the Identification of Indicative Linguistic Features. In McCarthy, P. M. and Boonthum, C. (Eds.), *Applied Natural Language Processing and Content Analysis: Identification, Investigation, and Resolution*. Hershey, PA: IGI Global.
- McNamara, D. S., and Graesser, A. C. in press. Coh-Metrix: An Automated Tool for Theoretical and Applied Natural Language Processing. In McCarthy, P. M. and Boonthum, C. (Eds.), *Applied Natural Language Processing and Content Analysis: Identification, Investigation, and Resolution*. Hershey, PA: IGI Global.
- Soykan, O. N. 2009. Arts and Languages: A Comparative Study. *Art Criticism* 24(1):113-121.