

An Analysis of the Gender Predictive Quality of Six Language Features in Freshman Essays

By Kim Fontana
Unknown Date

**Note — Please see MLK for related materials, such as tables, that are not included in the web version of this paper.*

RESEARCH REVIEW

In the nearly two decades since Thomas Farrell's ground-breaking article, "The Female and Male Modes of Rhetoric," appeared in *College English*, an enormous amount of scholarly interest and speculation has grown up around the subject of gender and language. Today there are several now famous analyses of the speech of women versus men as well as the style and content of their written discourse.¹ In these same years, acceptance of qualitative research as a legitimate alternative has grown tremendously. Perhaps as a result of these contemporaneous trends in language scholarship, there are virtually no studies which have truly quantifiable language features and gender as their focus. Ironically those researchers who have sought to measure specific language features in the writing of men and women have had to use data compiled for entirely different purposes. The extensive studies of Kellogg Hunt and W. Loban, designed to assess compositional maturity and development, have frequently provided the best databases for comparing specific language features of the writing of boys and girls. The lack of data in which gender and grammatical structures interact, coupled with the paucity of interest in quantitative research, has left a troubling gap in our understanding of the writing of men and women. The conspicuous absence of such inquiry is all the more regrettable given our continuing fascination with the subject of gender and language.

One recent study attempts to redress this gap in our understanding by connecting the tools of quantitative research with our interest in gender and written language. Anthony Mulac, Lisa Studley, and Sheridan Blau compiled a list of nineteen quantifiable language features which had been linked in some way to gender.² They obtained impromptu essays from students in the fourth, eighth, and twelfth grades and subjected these essays to regression analysis using the nineteen potential gender predictive features as the variables (Mulac, Studley, and Blau 447). This study obtained remarkable results. Analysis of the presence of these features in the essays resulted in 84 to 87% accuracy in gender prediction across the three grade levels.

The present study attempts to reproduce the results of Mulac et al.

Design

Of the original list of nineteen language features, the study by Mulac et al found four to be particularly indicative of male twelfth graders and two to be particularly indicative of female twelfth graders. The two features indicating female authorship are fillers and progressive verbs. Features indicative of male authorship are uncertainty verbs, subordinating conjunctions, judgmental phrases, and sentence initial adverbials. Mulac, Studley, and Blau define these markers (448-449). The present study adopts their study's markers and definitions and attempts to reproduce their results with a somewhat larger sample of the writing of slightly older students. The study's null hypothesis is that, by analyzing language features derived from the earlier study by Mulac et al, correct authorial gender will be predicted with no more than chance accuracy.

[\(Back to Top\)](#)

Method

Writing Samples

The kindness of Professors Ann Gebhard and Mary Kennedy provided copies of forty essays produced by freshmen at SUNY Cortland. Among the writers are twenty men and twenty women. All essays were typed and the copies were screened for obvious gender clues. In some cases, the attached assignment sheets included samples of student handwriting; these were removed. The essays are on various topics,

but all fall into the category of analytical writing. The length of the essays varies and was largely determined by their authors, nevertheless most are between 500 and 800 words.

Preliminary finding

As a preliminary procedure, the author read several essays and attempted to assign a gender to the writer on the basis of content. These gender ascriptions were mistaken more often than not. This finding confirmed other research that demonstrates that untrained readers (those operating without specific gender markers) cannot successfully differentiate between the writing of men and women (Mulac 442), although they may believe they can (Haswell and Tedesco 3).

Procedure

Sample length

To control for length, only each essay's first 200 words, last 200 words, and any words attached to these to make complete paragraphs were selected for analysis. In this way, each writing sample includes features potentially representative of the essay's introduction, body, and conclusion. The essays often include sizeable quoted passages. Quotations of less than one sentence are counted as part of the sample. Quotations of one full sentence or more are not counted in the sample. No quotations are analyzed for language features. As a result of this procedure, the samples analyzed range in length from 808 to 359 words. This fairly large range resulted from the sizeable differences in length of first and last paragraphs. In a few cases, the sample comprised the entire essay.

Language Features

The six language features Mulac found to be indicative of gender in twelfth graders are the following: subordinate conjunctions, initial adverbials, judgmental phrases, uncertainty verbs, fillers, and progressive verbs. This study analyzes the same features in these samples of the writing of college freshmen. Below are definitions and examples of each language feature.

Subordinate Conjunctions: words that join grammatically unlike elements and introduce dependent clauses whether adverbial, adjectival or nominal. This study makes use of a standard list of subordinate conjunctions including those pertaining to time, e.g., after, before, once, since, until, when, whenever, while; those pertaining to cause, e.g., as, because, since; those pertaining to result, e.g., in order that, so, so that, that; those pertaining to condition, e.g., if, even if, provided that, unless; those pertaining to contrast, e.g., although, even though, though, whereas; those pertaining to location, e.g., where, wherever; and those pertaining to choice, e.g., rather than, than, whether (Troyka 71). If subordinate conjunctions also met the criteria for initial adverbials (see below), they were counted as initial adverbials and not subordinate conjunctions.

Initial Adverbials: adverbial clauses which begin sentences. These clauses answer the questions, "how? when? or where?" with respect to the main clause and must appear at the beginning of the sentence. Typically sentences beginning with subordinate conjunctions, prepositional phrases, or often conjunctive adverbs, e.g., therefore and however, met the criteria for initial adverbials. Composition 86 provides two examples of sentence initial adverbials; the second of these is a prepositional phrase. "After reading *The Rich and the Poor* by Myron Magnet and *Why are they Homeless?* by Peter Marcuse, I became aware In the area of the deinstitutionalization of the mentally ill both Marcuse and Magnet have similar ideas (1). "After reading" and "In the area" were both considered sentence initial adverbials.

Uncertainty Verbs: those that indicate a lack of certainty on the part of the writer, e.g., "seems, may, could." The writer of composition 52 uses "seemed" in an expression of uncertainty: "One problem that has seemed to arise is that many schools are not equipped with the facilities needed for the proper education of the child" (1).

Progressive Verbs: those verbs presented in the "-ing" form appearing with some form of the verb to be. The writer of composition 24 uses "telling" and "thinking" as progressive verbs: "this sentence is telling the reader that no matter how much one has, a greedy person will still be thinking of something bigger and better to have" (1). Progressive verbs appeared in all tenses.

Judgmental Phrases: those statements which indicate personal evaluations rather than mere description. Composition #73 provides the following example: "Violence, however, is wrong and inexcusable no matter what form it comes in."

Fillers: these are words used for other than their semantic meaning. "Well" and "like" are often fillers. The following is a prototype: "Well, like if she was really smart." This study's essays offered no examples of

fillers.

[\(Back to Top\)](#)

PROCESS OF ANALYSIS

Incidence of language features

For each sample, the word count and incidence of each of the six language features were recorded. No examples of fillers were found in any essay sample. See table I below. Then an incidence per each 100 words was established by dividing the number of times a feature appeared by the sample's word count and then multiplying by 100. For example, the sample from essay #27 contained 595 words and 15 initial adverbials. Therefore, the initial adverbial incidence for essay #27 is 2.52 or $(15/595)100$. The incidence of each feature in each essay was, in this way, factored. These results are expressed as feature frequencies. See tables II-VI below. In addition, the average incidences for each feature were figured by averaging the word counts and the incidences of each feature and then using the foregoing equation. For instance, the average word count of the samples was 528 and the average incidence of initial adverbials was 5.7. Therefore, the average incidence of initial adverbials per 100 words was 1.08 or $(5.7/528)100$.

(See) Table I

Creation of coefficients expressing incidence relative to average

Once the average incidence for each feature was factored, individual essays' features could be compared using a feature coefficient. This coefficient expresses the incidence of that feature against the average by simply figuring the difference. For instance, essay #27 has an initial adverbial frequency of 2.52. Because the average frequency for initial adverbials is 1.08, essay #27's initial adverbial coefficient is 1.44. In this way, a coefficient was ascribed to each feature for every essay. These coefficients served as the numerical basis for predicting the gender of the essay's writer.

(See) Tables II-VI

Transformation of coefficients into an expression of gender

The features' coefficients were embellished with either a male or female appellation depending on whether, on the basis of the study by Mulac et al, we hypothesize the feature to be indicative of male or female writing. For instance, since essay #27 has a positive coefficient for initial adverbials (i.a.), and since we assume that feature is indicative of male writing, the coefficient for initial adverbials in essays #27 was altered to read 1.44m. Essay #47 has an ia coefficient of -.55 which indicates that it fell below the average incidence by that amount. The transformed initial adverbial coefficient for this essay becomes .55f. In this way, each coefficient was transformed into a positive number indicating either male or female quality. Each essay then received an M score (the sum of its male coefficients) and an F score (the sum of its female coefficients).

(See) Table VII

Gender Prediction

Finally, the sum of each essay's male coefficients was compared against the sum of each essay's female coefficients. The difference between these served as the basis for gender prediction. In this way, each essay acquired a specific score, labeled either male or female. For instance, essay #27 has a female coefficient of .634 and a male coefficient of 1.526; the essay's final, gender predictive score, was thus .892m. This gender predictive score formed the basis for the gender prediction for each writer.

(See) Table IX

RESULTS

The results of this study prove the validity of its null hypothesis: correct gender could not be ascribed on the basis of language feature analysis with more than chance accuracy. Final scores predicted twenty of the essays to be written by men and twenty by women. Nine of the twenty predictions of male authorship were correct, and nine of the predictions of female authorship proved correct. Of the twenty essays predicted to be written by female authors, nine were; of the twenty essays predicted to be written by men, nine were. 18 out of 40 predictions proved correct; gender was predicted with 45% accuracy.

[\(Back to Top\)](#)

ANALYSIS

This study failed to find six language features to be predictive of gender. The difference between the predictive accuracy of Mulac's study (84%) and the present study (45%) is startling. The most likely explanations of this discrepancy are differences in the writing samples and methods of analysis.

Regression analysis, the method the study by Mulac et al uses, is complicated and enables the researcher to weight the variables differently. This study's method treats each variable as equally significant. It seems likely that the nature of regression analysis was a factor in the differing results.

An equally plausible and more interesting explanation for the discrepancy in accuracy may lie in the differences between the writing samples. The writers used by Mulac et al were twelfth graders; essayists for the present study were college freshman. It is possible, although unlikely, that developmentally the writers were different and a different set of features would have been predictive of the genders of these slightly older students. Mulac, Studley, and Blau report that some features were found to be predictive of male writers at one grade level and female writers at another (451-52). Nevertheless, any developmental difference in the writers was slight and was far overshadowed by the probably more significant difference between the two writing samples--the writing task.

The essays Mulac, Studley, and Blau studied were impromptu drafts composed at one sitting. The students were prompted to write on the topic, "Is it important to tell the truth? Can it ever be better to lie?" This topic is similar to many the college freshmen chose to write on insofar as it invites analysis and argument. However, the writers for the study by Mulac et al were specifically advised to draw from their own experience. While some of this study's writers did draw from personal experience, they were not necessarily advised to do so. Such a difference in direction, may have resulted in differences in the appearance of a given feature. For instance, it is likely that the relative abundance of judgmental phrases that Mulac et al found, resulted from the nature of the topic and the teachers' request that the students use examples from their own lives.

The writers studied by Mulac, Studley and Blau produced a draft. They had no opportunity to revise. Essays for this study were drawn from process-oriented composition classes and consisted of final copies of essays which may have undergone many drafts. This difference may have affected the samples in such a way as to lessen the predictive quality of the features. Revision is surely responsible for the utter lack of fillers in the college sample. The college writers' revisions may have reduced the prevalence of uncertainty verbs and progressive verbs as well. In revision, students may have elected to replace relatively weak progressive verbs with stronger action verbs. Similarly, students may have chosen to replace tentative phrases employing verbs of uncertainty with more emphatic statements as their theses solidified. These effects of revision are speculative. If present, however, they significantly alter the sample and the relevance of these particular features to the sample. If the nature of the sample precluded fillers (as it seems it did) and discouraged progressive verbs (as it well may have), then the potential of this set of features to predict female authorship becomes virtually nil and the consequent potential accuracy of gender prediction is reduced by fifty percent.

Various studies have demonstrated that writing differs by task more than by other variables.³ While this study was not designed to assess writing differences by task, corroboration of this view may be its most important contribution. It seems clear that a given set of features, predictive of gender in one writing context, may not be predictive of gender in another writing context even if that context is basically similar.

[\(Back to Top\)](#)

NOTES

1 Among the most often cited analyses of style are Robin Lakoff's works, which include her *Language and Woman's Place* (New York: Harper and Row, 1975) and "Women's Language," *Language and Style* 10 (1977): 222-47, and M. Hiatt's "The Feminine Style: Theory and Fact," *College Composition and Communication* 29 (1978): 222-32. Studies more closely directed at the content of writing include: Catherine M. Lynch and Mary Strauss-Knoll's "Mauve Washers: Sex Differences in Freshman Writing," *English Journal* 76 (1976): 90-94, and Jo Keroes' "But What Do They Say? Gender and the Content of Student Writing," *Discourse Processes* 13 (1990): 243-57.

2 Anthony Mulac, Lisa Studley, and Sheridan Blau, "The Gender-Linked Language Effect in Primary and Secondary Students' Impromptu Essays," *Sex Roles* 23 (1990): 439-69. Most of the features Mulac and his associates chose were linked to speech. In several cases, the same feature had been linked to both sexes in different studies. In addition, Mulac's study included language features linked to speakers and writers regardless of age. Essentially, Mulac and his associates regarded any even casual connection

between a language feature and gender as potentially predictive of the gender of the writer.

3 See especially Donald L. Rubin and Kathryn Greene's "Gender-Typical Style in Written Language," *Research in the Teaching of English* 26 (1992): 7-39. This study compares lexical and syntactic features of essays produced under what the authors believe to be opposite ends of the compositional continuum. Half of the compositions were personal, spontaneous, and expressive; these were written for an intimate audience. The remaining compositions were revised, instrumental essays written for a distant audience. The researchers found that differences "due to mode of discourse were more widespread than differences due to gender" (7).

REFERENCES

Farrell, Thomas J. "The Female and Male Modes of Rhetoric." *College English* 40.8 (1979): 909-21.

Haswell, Richard and Janis E. Tedesco. "Gender and the Evaluation of Writing." Annual Meeting of the National Council of the Teachers of English. Seattle, WA 22-27 Nov. 1991.

Hiatt, M. "The Feminine Style: Theory and Fact." *College Composition and Communication*. 29 (1978): 222-32.

Hunt, Kellogg W. *Grammatical Structures Written at Three Grade Levels*. Champaign, IL: National Council of Teachers of English, 1965.

Keroes, Jo. "But What Do They Say? Gender and the Content of Student Writing." *Discourse Processes* 13 (1990): 243-57.

Lakoff, Robin. *Language and Woman's Place*. New York: Harper and Row, 1975. "Women's Language." *Language and Style* 10 (1977): 222-47.

Loban, W. *Language Development: Kindergarten Through Grade Twelve*. Urbana, IL: National council of Teachers of English, 1976.

Lynch, Catherine M. And Mary Strauss-Noll. "Mauve Washers: Sex Differences in Freshman Writing." *English Journal* 76 (1976): 90-94.

Mulac, Anthony, Studley, Lisa, and Sheridan Blau. "The Gender-Linked Language Effect." *Sex Roles* 23 (1990): 439-69.

Rubin, Donald L. and Kathryn Greene. "Gender-Typical Style in Written Language." *Research in the Teaching of English* 26 (1992): 7-39.

Troyka, Lynn Quitman. *Simon & Schuster Handbook for Writers*. 3rd ed. Englewood Cliffs: Prentice, 1993.

[\(Back to Top\)](#)