A Methodology for the Analysis of Narrative Accounting Disclosures
Author(s): Katherine Beal Frazier, Robert W. Ingram and B. Mack Tennyson
Source: *Journal of Accounting Research*, Vol. 22, No. 1 (Spring, 1984), pp. 318-331
Published by: Wiley on behalf of Accounting Research Center, Booth School of Business, University of Chicago
Stable URL: http://www.jstor.org/stable/2490713
Accessed: 27-10-2016 08:54 UTC

Journal of Accounting Research
Vol. 22 No. 1 Spring 1984
Printed in U.S.A.

A Methodology for the Analysis of Narrative Accounting Disclosures

KATHERINE BEAL FRAZIER,* ROBERT W. INGRAM,† AND B. MACK TENNYSON††

1. Introduction

Accounting research methods typically rely only on quantitative data. Narrative data in accounting reports, news announcements, and audit reports, etc., are at best crudely considered. More frequently they are ignored altogether. This is in spite of the fact that management analyses, footnotes, and explanations of monetary data are routinely included as integral and presumably important elements of financial reports. On the basis of this, we believe an evaluation of the narrative elements in accounting reports may lead to a fruitful extension of research concerned with the information content of accounting information.

In this paper, we introduce a methodology developed for the purpose of statistically evaluating narrative data. The methodology consists of a content analysis system, called *WORDS*, which was developed explicitly for this purpose. The functions and logic of this system are first described. We then provide an illustration of the *WORDS* analysis using annual report disclosures. Finally, we consider potential applications for *WORDS* in accounting research.

2. The *WORDS* Computer Analysis System

In this section we describe the logic behind the *WORDS* system and discuss how the system derives statistical relationships between and

* Assistant Professor, University of Colorado; † Associate Professor, University of Iowa; †† Assistant Professor, College of Charleston. This research was funded in part by the Peat, Marwick, Mitchell Research Opportunities in Auditing program. Views expressed here are those of the authors and do not necessarily represent those of Peat, Marwick, Mitchell. [Accepted for publication April 1983.]

among narrative words which are descriptive of central themes in the text.

The *WORDS* system is based on a word-frequency contiguity logic—that is, certain words in a narrative that occur, and cooccur, frequently are assumed to represent the content of the narrative (see Iker [1974*a*; 1974*b*]). Language, both written and spoken, is naturally redundant. A considerable portion of the redundancy results from words which have little to do with thematic content, such as prepositions and conjunctions. If these noncontent words are omitted from the analysis, redundancy in nouns, verbs, and modifiers (content words) will represent the central theme of a narrative. This associational structure, or internal contiguities between symbols, was advanced by Osgood [1959, pp. 44–45] as a means of identifying thematic content: "If ever there is any content analysis technique which has a defensible psychological rationale, it is the contingency method. It is anchored to the principles of association which were noted by Aristotle, elaborated by the British Empiricists, and made an integral part of most modern learning theories. On such grounds, it seems reasonable to assume that greater-than-chance contingencies of items in messages would be indicative of associations in thinking of a source."

WORDS focuses on individual words and their statistical interrelationships. The process is described by Iker [1974*a*, pp. 95–96] as follows:

> Our unit of information is the word itself. Dividing an interview into segments of time, i.e., one minute segments, we count the frequency ... with which each word appears in each time segment. Using these data, intercorrelations among words are obtained; operationally, these correlations represent the degree of ... association between words as they are observed across successive segments of interview time. Our assumption is that words which correlate highly with each other have much in common in defining topic or content reference while those with low correlations have little in common. We therefore factor-analyze this matrix of word intercorrelations to determine in a systematic fashion, if there are common factors which can account for the obtained correlations in an efficient and meaningful way.

While Iker speaks here of interviews (from psychoanalytic sessions), any division of text into segments (e.g., paragraphs, pages, cases) may be used.

The system which accomplishes this sophisticated series of maneuvers consists of over 30 modular programs. Callable in almost any order, these PL/1 programs allow two basic operations on any narrative. First, words in a narrative are subjected to a linguistic analysis consisting of frequency listings of content words. Second, the words enumerated in the first step are statistically analyzed. This requires the experimenter to order *WORDS* programs to carry out logically functions of grammatical analysis of the narrative and decide which statistical methods to apply.

In the linguistic analysis phase, narrative data must be entered in computer-readable form (card image, 80 character). Data may be entered with minimal editing. Contractions must be expanded to full form and proper nouns should be designated by preceding them with "$" signs. A code must be supplied for each observational unit (as part of the card image). Optional editing of the text is discussed later.

The system splits the text into words and tags each word according to its location in the text. A parsing program then deletes noncontent words and lemmatizes content words to root form. For example, different forms of the verb "like" ("liked," "likes," "liking") are considered to be equivalent in later frequency counts and intercorrelation measures. The system classifies each word according to a syntactic analysis logic developed in some detail by Klein and Simmons [1963]. The parts of speech of some common words are identified from an internal dictionary. The parsing program identifies the parts of speech of other words by means of morphological endings which indicate parts of speech. For example, "-ing" (verb), "-ed" (verb), and "-ly" (adverb) are included in the morphological classification scheme. Exceptions to these general morphological rules, such as "thing," "feed," and "sly," appear in the dictionary.

Though many words will receive a unique classification by either the dictionary or their morphologies, some words are not amenable to this type of unique classification. The first parsing pass performed by *WORDS* will assign a unique part of speech to each word where possible, but others will be assigned as many as three "most probable" parts of speech (for the most complicated cases). Then, using a second pass, *WORDS* uses the sentence context as a classification guide. Words are "framed" for the analysis. For example, in a sentence, "The ________ is in the sky," only a noun could properly fill the blank as the subject of the sentence. A particular word remains for subsequent analysis only to the extent that it fits the requirements of a sentence.

After parsing, *WORDS* sums the word frequencies and chooses up to 215 of the remaining words for statistical analysis. The 215-word limit is a current size limitation of the correlation program which defines the maximum size of the intercorrelation matrix. If the system is permitted to operate by default, the 215 most frequent words will be included in the remaining analysis.

Since words with high frequencies may not always be those of greatest interest, a subroutine of *WORDS*, *SELECT*, allows potentially more meaningful words to be extracted. The routine maximizes correlation coefficients, computing the absolute correlation between each word combination. The potential importance of a word is measured by the sum of its paired coefficients. Words with the highest correlation sums are retained for further statistical analysis. A screening technique (for correlations below some minimum size) may be used to remove chance correlations. An exponentiating routine prevents words with many small correlations from eliminating words with few high correlations. *SELECT* produces a subset of words with higher intercorrelations than those produced by frequency selection alone. Nevertheless, Iker [1975, p. 3] suggests: "whether the '*n*' highest frequency words or the '*n*' highest associating words are chosen, the same set of major themes should, in general, be identified."

At this stage of the analysis, *WORDS* has identified a subset of words from a text that is potentially useful for describing the thematic content of the data. Each word is identified by a code, and the attributes of the word that make it unique or similar to other words in the text are identified. Further analysis can be either by factor analysis, using the Kaiser principle-component algorithm followed by varimax rotation to a simple structure, or by a recursive method producing higher and higher order word clusters. (Both are derived from the intercorrelation matrix.) Iker [1975] argues that both methods produce essentially the same "themes" or word groups.

If a factor analysis is used, factor scores are produced showing both the segments of text in which each word factor occurs (and the strength of occurrence) and how factors are related for each segment. The factor scores are usable in subsequent analysis as measures of thematic content.

WORDS defaults to certain "reasonable" measures throughout, though most of these defaults can be overridden at the user's option. For example, when calculating the intercorrelation matrix, two words are observed as "associated" in an occurrence when they are within five words of one another. The distance can be expanded or contracted by a simple default override. The default of five was set to correspond with typical phrase and sentence length.

Although *WORDS* does not explicitly join words which are synonyms, synonyms can be joined and analyzed as a single variable using editing routines. Typically, the time and subjectivity involved in the synonymization process prohibit its use. Harway and Iker [1965] experimented with synonymization, but concluded that content themes derived with the process are so similar that they do not alter the interpretation of output.

Finally, *WORDS* has editing routines by which the experimenter can edit in the words wanted for subsequent analysis or can edit out unwanted words. Thus, after examining parsed lists of words the experimenter may choose to use any relevant portion of the initial body of words by editing those words for the statistical analysis phase. For example, special vocabulary typical of formal accounting disclosures and audit reports can be identified for special consideration using editing programs.

Appendix A contains additional technical information about the *WORDS* system. Availability and cost of the system are also discussed in Appendix A.

3. An Illustration

WORDS can be applied to any narrative data. In the following illustration, we applied *WORDS* to Management Analyses of the Results of Operations from the 1978 annual reports of 74 firms.[1]

[1] These firms were from the metal mining and manufacturing, oil, and chemical industries that were used in a previous study.

The sample was initially divided into positive and negative performance and whether Management Controlled (*MC*) or Owner Controlled (*OC*) groups. A firm was assigned to the *OC* group if at least 10% of the voting stock was held by one party.[2] Seventeen of the 74 firms in the sample met the *OC* criterion. The measure of performance used to group firms into positive and negative performers was based on the sign of the difference between the firm's percentage earnings growth from 1977 to 1978 relative to the percentage earnings growth from 1976 to 1977. Firms that demonstrated (lower) higher growth in 1978 than in 1977 were classified into the positive (+) (negative (−)) performance group, using income before extraordinary items.[3]

The factor scores from the *WORDS* analysis were employed in a two-factor multivariate analysis of variance test. The independent variables of interest were the control factor (*MC* or *OC*) and the performance factor (+ or −). In addition to the two main effects, we also examined interaction between them.

The purpose of our tests was to examine differences in annual report disclosures of firms grouped according to hypothesized incentives for management to misrepresent firm performance. Recent research has shown that good news (positive performance) is presented both in a more understandable form (Morton [1974] and Adelberg [1979]) and sooner (Givoly and Palmon [1982] and Patell and Wolfson [1982]) than bad news (negative performance). Other research has shown that *MC* firms adopt certain accounting procedures (Dhaliwal, Salamon, and Smith [1982]) and make accounting policy changes (Salamon and Smith [1979]) that are different from *OC* firms. These actions were interpreted to be an effort to misrepresent firm performance on the part of *MC* firms. Hence we hypothesized that *MC* firms would be motivated to misrepresent narrative reports, relative to *OC* firms.

In addition, we tested the predictive ability of the disclosures using factor scores from the *WORDS* analysis as independent variables in discriminant models. The sign of the cumulative average residual (*CAR*) for 1979, the year following the annual report year, was the dependent variable. Classification accuracy was tested for + and − performance, *MC* and *OC* firms.

Factors derived from the *WORDS* analysis are presented in table 1. Twelve factors with eigenvalues in excess of 1.0 are described there, with each consisting of a series of intercorrelated words listed in order by loading. The words were obtained first by limiting the analysis to high frequency count, content words. That is, words that did not appear in

---

[2] Other criteria could be employed. However, the effects of these changes on the present sample were considered to be minimal since few borderline cases existed.

[3] Other performance models also were evaluated. These included a cross-sectional earnings model and an abnormal security return model. Only minor differences were observed among these alternatives.

TABLE 1

Summary of Factor Word Loadings

Factor	1	2	3	4	5	6	7	8	9	10	11	12
Eigenvalue	23.5	14.8	6.9	3.8	3.6	2.9	2.9	2.7	2.5	2.3	2.0	1.9
Percentage of Explained Variance	7.36	8.39	3.41	2.78	4.07	2.64	2.20	2.15	2.08	2.89	2.40	1.94
	Future	Tax	Debt	Require	Record	Compare	Strike	Four	Raw	Common	Volume	Return
	Position	Segment	Fiscal	Environment	Domestic	Earn	Coal	Quarter	Material	Dividend	Revenue	Invest
	Strong	Lose	Up	Capital	Decline	Equity	Result	Substantial	Profit	Stock	Decrease	Government
	Good	Sell	Acquire	Finance	Shipments	Operate	Income	Major	Cost	Shareholder	Attribute	Current
	Continue	Expense	Stock	Construct	Gain	Substantial	Shipments	Improve		Share	Price	Capital
	Progress	¢*	Long-term	Control	Demand	Last	Effect	Earn		Cash	Increase	Regulate
	Grow	Cost	Cash	Improve	Foreign	Year	Demand			Up		Economy
			Affect	Facility	Customer	Lose				Rose		Improve
				Cost	Grow					Strong		

* Indicates monetary value.

the combined texts at least ten times and noncontent words (articles, prepositions, etc.) were omitted from the analysis.

Each of the factors was derived objectively from the text. Unfortunately, interpreting the factors is a subjective process. Most of the factors shown reflect intuitively obvious concepts. More precise relationships among the words were determined by reference to the texts of firms that scored high on each factor. A summary of the primary theme reflected by each factor is provided in table 2.

The themes suggest a variety of events that indicate performance characteristics of firms or explanations for their performances. A majority of the themes are related to events or activities that potentially increased costs and liabilities or decreased profits. Factors 2, 3, 4, 5, 7, 9, 11, and 12 are all of this general type. The other factors suggest improved or improving performance. Negative performance attributes are viewed in terms of external (generally uncontrollable) causes. (See Frazier [1982] for a more formal test of this phenomenon.)

Factor scores for each firm for each factor were used as dependent variables in the analysis of variance tests. These scores measure the importance of each factor for each firm. Mean scores for the analysis groups are presented in table 3. These means provide relative measures of the magnitude of the factors within each group.

Examination of table 3 reveals several aggregate differences across groups. For example, relatively higher factor scores were observed for theme 1 in the positive return groups both for *MC* and *OC* firms. This relationship is logical since theme 1 concerns progress and growth. Thus, table 3 provides a means for assessing the relationships between the themes and groups for those factors that are found to be significant.

The significance of each of the relationships was measured by the F scores reported in table 4. Three of the 12 themes—2, 5, and 6—were significant at $\alpha = .10$. Theme 2 described tax effects on segment losses;

TABLE 2
Summary of Factor Themes

Factor	Theme
1	Continued progress, strength of future position
2	Tax effect of segment losses
3	Increase in debt to reduce stock
4	Environmental improvements
5	Domestic declines accompanied by foreign growth
6	Comparison of earnings to last-year losses
7	Effect of strike on demand and income
8	Substantial improvement in fourth-quarter earnings
9	Effect of raw material costs on profits
10	Increase in common dividends
11	Decrease in revenue resulting from price increase
12	Effect of government regulation on return on investment and expectation for improvement in economy

TABLE 3

Factor Means for Management Analysis Themes

Theme	Manager Controlled		Owner Controlled	
	Negative Performance ($n = 37$)	Positive Performance ($n = 20$)	Negative Performance ($n = 11$)	Positive Performance ($n = 6$)
1	−.034	.309	.028	.457
2	−.177	.895	.033	.315
3	−.173	−.192	−.223	.413
4	−.141	.127	−.113	.355
5	−.120	.082	.189	−.433
6	−.045	−.451	−.272	.507
7	−.119	−.429	−.282	.063
8	−.157	.422	.069	.203
9	.075	.131	−.364	−.545
10	−.130	.128	.046	.003
11	.031	−.205	.672	−.165
12	−.128	.666	.042	−.090

TABLE 4

MANOVA Statistics

Theme	Model			Control		Performance		Interaction	
	F	P	R^2	F	P	F	P	F	P
1	.72	.55	.03	.15	.70	2.00	.16	.02	.89
2	3.30	*.03*	.134	.01	.92	8.59	*.00*	1.30	.26
3	2.00	.12	.086	1.40	.24	.94	.34	3.65	.06
4	.89	.45	.040	.24	.62	2.26	.14	.17	.68
5	2.94	*.04*	.121	.01	.94	.00	.99	8.80	*.00*
6	2.64	*.06*	.110	.81	.37	.34	.56	6.77	*.01*
7	.87	.47	.039	.09	.76	.54	.46	1.97	.17
8	1.96	.13	.084	.14	.71	4.92	.03	.82	.37
9	1.15	.34	.051	3.29	.07	.00	.99	.16	.69
10	.33	.80	.015	.08	.77	.61	.44	.31	.58
11	.99	.41	.044	1.13	.29	1.26	.27	.57	.45
12	1.89	.14	.082	.18	.67	3.67	.06	1.83	.18
Multivariate effects				1.06	.42	1.51	.12	2.01	.03

$F = F$ score.
P = Probability associated with F.

theme 5 discussed domestic declines and foreign growth; and theme 6 compared current earnings with last year's losses. Further examination of the sources of these results indicated that theme 2 was a return effect, whereas themes 5 and 6 were interaction effects.

Table 3 reveals that theme 2 was associated with positive return groups for both the *MC* and *OC* firms. The tax effects of segment losses apparently were of more importance for firms with positive performances than negative performances. This relationship is logical since cash-flow benefits of segment losses are more immediate for profitable than for unprofitable firms.

Themes 5 and 6 reveal differences between *MC* and *OC* firms when considered in relationship to firm performance. Table 3 shows that theme 5 was most important for the −*OC* and +*MC* firms. Theme 6 was most important for the +*OC* firms. The nature and direction of these associations do not provide clear evidence of a misrepresentation effect. Thus, while the nature of the themes derived from the texts were amenable to the possibility of management attempting to alter its perceived performance, the evidence obtained does not indicate that *MC* firms were more inclined to this activity than *OC* firms or that "−" performance firms were more inclined than "+" performance firms.

While none of the remaining *ANOVA* models was significant, several additional effects in some of these models were. Even if one were to posit that these differences were important, the conclusions would not change. The control effect in theme 9 reveals that *MC* firms discussed raw material costs more than *OC* firms, but this was true for both + and − groups. + firms discussed theme 8, improvement in fourth-quarter earnings, more than − firms for both *MC* and *OC* groups. +*OC* firms were most associated with increases in debt to reduce stock.

The results of these tests do not support the hypothesis. The thematic content identified by *WORDS* appears to be representative of disclosures that are readily observable in annual reports. The identified relationships between the independent and dependent variables are logical and consistent with respect to explaining firm performance. However, the themes did not distinguish clearly between *OC* and *MC* firms, except in a few cases. More thematic similarity than difference was observed.

These results may indicate that the narratives concentrate on a number of environmental attributes that are common across firms rather than concentrating on individual performance; or the *WORDS* analysis may have been more sensitive to these themes. Alternatively, bad news firms (poor performers) may provide signals that imitate good news firms, thus interjecting ambiguity into the disclosures. A third possibility is that the disclosures are ex ante rather than ex post signals of performance, as posited by Copeland [1978]. The latter two alternatives can be tested by examining the predictive ability of the disclosures.

In line with this, we examined the predictive ability of the annual report content scores. The *WORDS* factor scores were used in a discriminant model to evaluate their association with the sign of the *CARs* for 1979. If managers of − performance firms were attempting to misrepresent performance, the disclosures provided in the annual reports should not have been indicative of the firms' future performances. This analysis may also provide a test of Copeland's [1978] hypothesis concerning the information value of unaudited annual report disclosures.

Both *MC* and *OC* firms were combined in this analysis since few differences were observed previously. In addition, the differences in

TABLE 5
Percentage of Correct Classification for Discriminant Models

1979 Return	Nonholdout Model			Lachenbruch Model		
	MC (n=57)	*OC* (n=17)	Total (n=74)	*MC* (n=57)	*OC* (n=17)	Total (n=74)
−*CAR*	75.0	83.3	76.7	70.8	83.3	73.3
+*CAR*	72.7	81.2	75.0	66.7	72.7	68.2
Total	73.7	82.4	75.7	68.4	76.5	70.3

sample sizes between groups could bias the results in favor of the *OC* firms since the ratio of variables to observations would be greater.

Covariance matrices of the +*CAR* or −*CAR* groups were compared and found to be unequal. Accordingly, a quadratic model was employed. The Lachenbruch (Jackknife) procedure was used to determine the predictive ability of the *WORDS* factor scores.[4] Monthly abnormal returns were computed from the simple market model using the *CRSP* value-weighted, dividend-adjusted index as a proxy for the market index. Model parameters were computed using the 72 monthly returns from the 1973–78 fiscal years for each firm. The *CAR* for each firm was computed as the sum of the monthly abnormal returns for fiscal 1979.

Table 5 reveals the correct classification results for both a nonholdout model and for the Lachenbruch model. The small amount of deterioration in classification ability of these models indicates that the model parameters are relatively insensitive to changes in the sample.

Correct classification rates for the entire sample were better than 70% for both −*CAR* and +*CAR* groups. This rate is significantly better than chance at $\alpha = .05$. The *OC* classification rates were somewhat higher than the *MC* rates; however, the differences can be accounted for by chance. The findings are consistent with the hypothesis that management analysis data in the annual report are useful for predicting the future performance of a firm.

4. *Accounting Applications of WORDS*

Narratives in annual reports were chosen for preliminary studies using words. For example, Frazier [1981; 1982] compared annual report narratives with financial performance attributes. Ingram and Frazier [1983] described a number of financial and economic factors associated with narrative disclosures. Tennyson [1982] examined the relationship between narrative disclosures and bankruptcy. Potential applications are wide-ranging. *WORDS* has been successfully applied in many contexts (see Harway and Iker [1969], Iker [1974; 1975], Klein and Iker [1974], and Oakman [1980] for descriptions). It has proved to be a robust technique across widely varied research problems.

[4] This procedure permits the use of a small sample by iteratively classifying each observation based on a model derived from the remaining $n - 1$ observations.

Currently, *WORDS* is being adapted to aid in a verbal protocol study. Methodologically, *WORDS* may become a coding scheme which enables the researcher to avoid some of the problems articulated by Libby [1981, pp. 94–95]:

> Questions concerning the objectivity of data-coding methods, in particular those related to verbal protocols, are often so severe as to question the scientific status of the research. In fact, practitioners of the technique admit that this portion of protocol analysis is more of an art than a science.... Even the most objective computer-aided coding techniques require only that the coder (usually the researcher or assistant) be able to force the protocols into a predetermined general framework. This can lead to selective attention to certain protocols and lack of reproducibility of results. Furthermore, alternative coding schemes which could as easily "fit" the data are usually readily available. Attempts to measure agreement between different coders using the same scheme do not effectively address this issue. A comparison with competing coding schemes more closely approaches a solution to this problem.

Since *WORDS* does not depend on a predetermined framework or subjectively derived coding categories, the results should be reproducible. Data reduction using *WORDS* could introduce interpretive clarity otherwise unavailable in lengthy, convoluted protocols. Indeed, in its initial application, *WORDS* was used on transcripts of psychoanalytic sessions of a subject during hundreds of therapy hours (see Harway and Iker [1965]). Results were clear and easily interpreted. The close parallels to verbal protocols are encouraging.

Self-reports of decision processes are only one of the many potential uses for *WORDS*. In a study by Klein and Iker [1974], "The Lack of Differentiation between Male and Female in Schrebner's Autobiography," contextual correlations of the words "male" and "female" were virtually indistinguishable. Parallel applications in accounting might indicate contexts which allow the reader to distinguish, or not, between firms' different conditions. For example, qualified audit opinions presumably allow the reader to distinguish among different audit results. However, if contexts of the qualifications are basically invariant across firms, the user's ability to discriminate may be impaired.

Similarly, potentially useful data for event studies could be adapted to analyze and classify events such as news announcements. Classified narratives could then be entered along with other quantitative data for subsequent analysis. The extent to which even subtle elements of language choice and use affect perceptions and reactions to announcements could be considered. Many news announcements (other than preliminary earnings announcements), such as those in the *Wall Street Journal*, consist of narrative discussions. Pertinent events data must be abstracted from these narratives and coded in a fashion suitable for statistical analysis. Examples of accounting event studies that use coded news announcements include Dopuch, Holthausen, and Leftwich [1983] and Waymire [1983].

Other narrative elements in annual reports could be analyzed with

WORDS. Footnote studies such as those by Morton [1974] and Adelberg [1979] indicate the fruitfulness of considering qualitative attributes of narrative disclosure and the related impact on a user's ability to understand and use the data. *WORDS* appears to be potentially useful for accomplishing an objective, systematic classification and quantification of typically difficult-to-study data.

APPENDIX A

A Technical Description of WORDS System Operations

When data are input, each word and punctuation in the text is separated into an independent record by the system. The standard record is broken into ten discrete fields, as follows (Iker and Klein [1974, p. 431]):

WORD. Contains the word that is the basic system datum. The field accommodates up to 16 characters.

INTV. A three-digit field indicating the interview (observation) in which the word was located.

SEGM. A three-digit number which identifies the segment of the interview in which the word was located (optional).

SEQ. A five-digit number indicating the sequence of the word in the segment (optional).

GTAG. A one-character field set (by the parsing programs) to designate the word's major part of speech.

FREQ. A five-digit field set at 00001 at input to indicate the word's frequency of occurrence.

SPKR. A single-character field designating the speaker who emitted the word (optional).

*AUX*1. A three-digit auxiliary field open to various uses depending upon the programs called.

*AUX*2. A one-character field, initially blank, reset by the parsing programs for ancillary *GTAG* information.

*AUX*3. A one-character field open to use by various programs. It is analogous to *AUX*1.

The successful use of the *WORDS* system depends on the correct configuration of a series of program calls to produce the desired kinds of output. Great flexibility is allowed by the system in controlling the individual programs used and the order of their use. Each program in the system has a name which is used to call it in the configuration. Iker and Klein [1974, pp. 432–33] describe briefly the use and function of each of the independent programs.

Because the system is so flexible, the input/output (*I/O*) manipulations, which pass data into, out of, and between programs, are very complex. Executive systems handle the *I/O* and schedule jobs. On IBM/370 machines the general executive is the operating system (*OS*). The language which directs the operating system is Job Control Language (*JCL*). A *JOBS* program allows the flexible use of *WORDS* programs without writing special *JCL* commands for every variation.

When analyzing data with *WORDS*, two computer runs are necessary. The control cards submitted are actually to *JOBS*, which in turn uses the deck to issue *JCL* statements which communicate with the *OS*. The second run is *WORDS*, actually prepared by *JOBS* rather than the user.

In *JOBS*, the programs desired are called by letter description. The output of any program can be passed on to any subsequent programs as input. Some of the functions handled internally by *JOBS* include: "(1) Describing the file; (2) computing and maintaining necessary space requests for the file; (3) controlling block information for the file; and (4) deleting (unless the user indicates otherwise) the file from the system as soon as it is determined that it will no longer be used again in the run" (Iker and Klein [1974, p. 430]).

While *WORDS* programs allow total flexibility in the order called, many programs are logically and almost inevitably called in sequence. The *JOBS* program allows "procedure" calls to be made which call a fixed sequence of independent programs with one call.

Diagnostic messages are also produced by *JOBS*. Though eventually *JOBS* will abort runs containing errors, it usually scans all the input data for additional error messages before issuing the command to abort. Over 100 diagnostic checks are contained in *JOBS*.

The ten fixed fields of the standard record are usable interchangeably in sorting to determine output order. Ordering the field(s) containing the data is specified by the user. For example, a new file in which all words in the data set are alphabetized can be produced. By sorting on "*INTV*," the new file would contain sorted words from interview (observation) 1, then interview 2, etc. Multiple parameters can be specified. Using this sorting flexibility, it is possible to obtain a list of word-type frequencies, to enumerate types for each interview, to list adjectives in the data, ordered frequencies, etc.

WORDS is available for the cost of the computer tape and copying time from Professor Howard Iker, School of Medicine, Rochester University, Rochester, New York.

While the acquisition cost is nominal, start-up and operating costs are not. The system is sensitive to operating characteristics of an installation and must be modified accordingly. At present, *WORDS* is most easily adapted to IBM-software-compatible installations. The system requires its own disk pack, and data analysis of large volumes of text requires a minute or more of CPU time. Once the data have been read into the system and are reformatted, subsequent analysis time can be reduced significantly.

REFERENCES

ADELBERG, A. "A Methodology for Measuring the Understandability of Financial Report Messages." *Journal of Accounting Research* (Autumn 1979): 565–92.

COPELAND, T. "Efficient Capital Markets: Evidence and Implications for Financial Reporting." *Journal of Accounting, Auditing and Finance* (Fall 1978): 33–48.

DHALIWAL, D., G. SALAMON, AND E. SMITH. "The Effect of Owner versus Management Control on the Choice of Accounting Methods." *Journal of Accounting and Economics* 4 (1982): 41–53.

DOPUCH, N., R. HOLTHAUSEN, AND R. LEFTWICH. "Selectivity in Financial Reporting: The Case of Qualified Audit Opinions." Working paper, University of Chicago, January 1983.

FRAZIER, K. B. "Word Frequency Contiguities in Narrative Accounting Disclosures." Ph.D. dissertation, University of South Carolina, 1981.

———. "Information Inductance and Management Attribution of Financial Results." In *Proceedings of the American Institute for Decision Sciences*. San Francisco, 1982.

GIVOLY, D., AND D. PALMON. "Timeliness of Annual Earnings Announcements: Some Empirical Evidence." *The Accounting Review* (July 1982): 185–222.

HARWAY, N., AND H. IKER. "A Computer Approach Towards the Analysis of Content." *Behavior Research Methods and Instrumentation* 6 (1974): 430–38.

IKER, H. P. "An Historical Note on the Use of Word-Frequency Contiguities in Content Analysis." *Computers and the Humanities* 8 (1974*a*): 93–98.

———. "SELECT: A Computer Program to Identify Associationally Rich Words for Content Analysis. I. Statistical Results." *Computers and the Humanities* 8 (1974*b*): 313–19.

———. "SELECT: A Computer Program to Identify Associationally Rich Words for Content Analysis. II. Substantive Results." *Computers and the Humanities* 9 (1975): 3–12.

———, AND R. KLEIN. "WORDS: A Computer System for the Analysis of Content." *Behavior Research Methods and Instrumentation* 6 (1974): 430–38.

INGRAM, R. W., AND K. B. FRAZIER. "Environmental Performance and Corporate Disclosure." *Journal of Accounting Research* (Autumn 1980): 49–60.

———. "Narrative Disclosures in Annual Reports." *Journal of Business Research* (March 1983): 49–60.

KLEIN, R. H., AND H. P. IKER. "The Lack of Differentiation between Male and Female in Schrebner's Autobiography." *Journal of Abnormal Psychology* (1974): 234–39.

KLEIN, S., AND R. F. SIMMONS. "A Computational Approach to Grammatical Coding of English Words." *Journal of the Association for Computing Machinery* 10 (1963): 334–47.

LIBBY, R. *Accounting and Human Information Processing: Theory and Applications.* Englewood Cliffs, N.J.: Prentice-Hall, 1981.

MORTON, J. R. "Qualitative Objectives of Financial Accounting: A Comment on Relevance and Understandibility." *Journal of Accounting Research* (Autumn 1974): 288–98.

OAKMAN, R. L. "Computerized Content Analysis of Narrative Annual Reports: A Computer in the Humanities Look at the Corporate Word." Paper presented at the International Congress on Applied Systems Research and Cybernetics, Acapulco, December 1980.

OSGOOD, C. E. "The Representation Model and Relevant Research Methods." In *Trends in Content Analysis*, edited by I. de Sala Pool. Urbana: University of Illinois Press, 1959.

PATELL, J., AND M. WOLFSON. "Good News, Bad News, and the Intraday Timing of Corporate Disclosures." *The Accounting Review* (July 1982): 509–27.

SALAMON, G. D., AND E. D. SMITH. "Corporate Control and Managerial Misrepresentation of Firm Performance." *Bell Journal of Economics* (Spring 1979): 319–28.

TENNYSON, B. M. "The Usefulness of Narrative Disclosure in Predicting Bankruptcy." Ph.D. dissertation, University of South Carolina, 1982.

WAYMIRE, G. "Market Resolution of the Moral Hazard Problem Associated with Corporate Voluntary Disclosures: Empirical Evidence on Managerial Forecasts of Annual Earnings per Share." Working paper, University of Chicago, January 1983.