On the role of computational, statistical, and interpretive techniques in multi-dimensional analyses of register variation: A reply to Watson‡

DOUGLAS BIBER

Abstract

The present paper addresses several issues relating to the research goals and methodological techniques used in multi-dimensional analyses of register variation (e.g., Biber, 1988). The paper is prompted by the allegations of inadequacy and error published in Watson (1994). In my discussion, I point out that Watson fails to substantiate his claimed inadequacies with empirical evidence, and that previous investigations have shown that many of these criticisms are factually incorrect. In addition, I attempt to correct several fundamental misunderstandings of the multi-dimensional approach and the relationships among computational, statistical, and interpretive methodologies within the approach.

Keywords *multi-dimensional analysis; register variation; corpus linguistics; computer taggers; multivariate statistics.*

1. Introduction*

The present paper addresses several issues relating to the research goals and analytical techniques used in multi-dimensional analyses of register variation.[1] The paper is prompted by the recent critique of the multi-dimensional approach published in this journal by Watson (1994): the allegations of inadequacy and error in that paper, which are largely unsubstantiated, are of such a serious nature that they require a direct response.

Judging from the title—'A multidimensional analysis of style in Mudrooroo Nyoongah's prose works'—the primary goal of Watson's paper is to apply the multi-dimensional analytical framework to an

‡ A reply to Gregory J. Watson, A multi-dimensional analysis of style in Mudrooroo Nyoongah's prose works, *Text* 14: 239–285 (1994).

0165–4888/95/0015–0341

analysis of author's style. Watson does undertake such an analysis, using the six major dimensions identified and interpreted in Biber (1988) to describe the stylistic characteristics of five novels written by Mudrooroo Nyoongah. By applying the multi-dimensional framework, Watson is able to identify and describe interesting developmental shifts in writing style over the course of this author's writing career.

However, Watson also sets a second goal for himself: 'to critically analyze and assess the usefulness of this model' (1994: 239, referring to the 1988 multi-dimensional analysis of English).[2] Watson's critique has three main points, as summarized in the following quotation:

> Biber's model (i.e., Biber, 1988) suffers from several serious inadequacies. These include his assumed connection between the communicative function of particular linguistic features and his qualitatively determined dimensions of style; his tagging procedure; and the size and choice of his original corpus.
>
> (1994: 280)

I address each of these three points in turn below. In my discussion, I point out that the claimed inadequacies cited by Watson are not substantiated by empirical evidence: Watson presents no empirical investigations of the computational techniques for grammatical tagging, and no empirical investigations of the representativeness of the corpus. In addition, it will be shown that Watson's criticisms reflect fundamental misunderstandings of the multi-dimensional approach and the relationships among computational, statistical, and interpretive methodologies within the approach.

Given Watson's near total absence of supporting evidence, coupled with obvious basic misunderstandings of multi-dimensional analytical techniques, it is distressing and disappointing that his repeated, detailed assertions of inadequacy and error could have found their way into print in a reputable, refereed journal. Such irresponsible claims are particularly culpable in that numerous published studies show that Watson's assertions are simply not accurate. In the following pages I hope to correct these misstatements.

In my discussion I first briefly introduce the multi-dimensional approach to register variation, focussing on the underlying research goals and the methodologies used for analysis. Then I discuss in turn each of the three major areas of criticism raised by Watson.

2. Overview of the multi-dimensional approach to register variation

The multi-dimensional approach to register variation was originally developed for comparative analyses of spoken and written registers in

English (e.g., Biber, 1986, 1988). Methodologically, the approach uses computer-based text corpora, computational tools to identify linguistic features in texts, and multivariate statistical techniques to analyze the co-occurrence relations among linguistic features, thereby identifying underlying dimensions of variation in a language.

The primary research goal of the multi-dimensional approach is to provide comprehensive descriptions of the patterns of register variation, including (1) identification of the underlying linguistic parameters, or dimensions, of variation, and (2) specification of the linguistic similarities and differences among registers with respect to those dimensions. Two primary motivations for the multi-dimensional approach are the assumptions that: (1) generalizations concerning register variation in a language must be based on analysis of the full range of spoken and written registers; and (2) no single linguistic parameter is adequate in itself to capture the range of similarities and differences among spoken and written registers. The approach thus requires analysis of numerous spoken and written registers with respect to numerous linguistic features.

Some of the general characteristics of the multi-dimensional approach are:

1. It is corpus-based, depending on analysis of a large collection of naturally-occurring texts.
2. It uses automated computational techniques to analyse linguistic features in texts. This characteristic enables distributional analysis of many linguistic features across many texts and text varieties.
3. It uses interactive computational techniques to check the analysis of ambiguous linguistic features, ensuring accuracy in the final feature counts.
4. The research goal of the approach is the linguistic analysis of texts, registers, and text types, rather than analysis of individual linguistic constructions.
5. The approach is explicitly multi-dimensional. That is, it is assumed that multiple parameters of variation will be operative in any discourse domain.
6. The approach is quantitative. Analyses are based on frequency counts of linguistic features, describing the relative distributions of features across texts. Multivariate statistical techniques are used to identify co-occurrence patterns among linguistic features and to analyze the relations among texts.
7. The approach synthesizes quantitative and functional methodological techniques. That is, the quantitative statistical analyses are interpreted in functional terms, to determine the underlying communicative

functions associated with each empirically determined set of co-occurring linguistic features. The approach is based on the assumption that statistical co-occurrence patterns reflect underlying shared communicative functions.

Dimensions represent distinct groupings of linguistic features that have been empirically determined to co-occur with significant frequencies in texts. It is important to note that the co-occurrence patterns underlying dimensions are identified quantitatively (by a statistical factor analysis) and not on any a priori basis. Dimensions are subsequently interpreted in terms of the communicative functions shared by the co-occurring features. Interpretive labels are posited for each dimension, such as 'Involved versus Informational Production' and 'Narrative versus Non-narrative Concerns'.

In earlier synchronic multi-dimensional analyses of English (e.g., Biber 1986, 1988), approximately 500 texts from 23 registers were analyzed, including face-to-face conversations, interviews, public speeches, broadcasts, letters, press reportage, official documents, academic prose, and fiction. Subsequent analyses have used this approach to analyze texts from a number of more specialized registers, such as elementary school textbooks and student writing, job interviews, and the writings of individual authors. Linguistic features analyzed in these studies include both lexical and grammatical characteristics of texts (see section 2.3 below).

Individual texts, or groups of texts called registers, can be compared along each dimension. Two registers are similar along a dimension to the extent that they use the co-occurring features of the dimension in similar ways. Multi-dimensional analyses show that a pair of registers are often similar along one dimension (i.e., with respect to one set of co-occurring linguistic features) but quite different along another dimension (i.e., with respect to another set of features).

2.1. *The use of automated and interactive computational techniques in multi-dimensional analyses*

The use of automated and semi-automated (i.e., interactive) computational techniques is a practical rather than necessary aspect of multi-dimensional analyses. Such analyses by hand would be extremely time-consuming, and they are often considerably less reliable and accurate than analyses by computer.

Before the use of computers, empirical discourse analyses were typically based on a few thousand words of text; an analysis of 10,000 words was regarded as a major undertaking that required a long research period.

Similarly it was possible to consider only a relatively restricted range of linguistic characteristics; analyses considering 10 different linguistic characteristics were regarded as major projects. In contrast, early multi-dimensional analyses employing computational techniques were based on a much more adequate and representative database: a text corpus over 100 times as large as in most previous analyses (nearly 1 million words of text), and inclusion of a very wide range of linguistic characteristics (67 different features in Biber [1988]).

Needless to say, some linguistic analyses must be checked interactively, because current automated techniques are not sufficiently accurate. For example, the distinction between some past tense verbs and past participial verbs functioning as post-nominal modifiers is notoriously hard for automated computer analyses. All automated grammatical taggers have difficulties dealing with distinctions such as this, and as a result, it is necessary to include interactive post-editing to insure accuracy (see section 3.2.2 below).

2.2. *Methodological overview of the multi-dimensional approach*

The multi-dimensional approach involves the following methodological steps:

1. Texts are collected, transcribed (in the case of spoken texts), and input into the computer. The situational characteristics of each spoken and written register are noted during data collection.
2. The published literature is reviewed, and if necessary supplemented by original grammatical research, to determine the range of linguistic features to be included in the analysis, together with functional associations of individual features (see, for example, Aijmer, 1984; Altenberg, 1984; Beaman, 1984; Chafe, 1982; Coates, 1983; Schiffrin, 1981, 1987; Tannen, 1982; Thompson, 1983; Tottie, 1986).
3. Computer programs are developed for automated grammatical analysis, to 'tag' all relevant linguistic features in texts.
4. The entire corpus of texts is tagged automatically by computer.
5. All texts are post-edited interactively to insure that the linguistic features are accurately identified.
6. Additional computer programs are developed and run to compute frequency counts of each linguistic feature in each text of the corpus.
7. The co-occurrence patterns among linguistic features (across all texts in the corpus) are analyzed, using a factor analysis of the frequency counts.

8. The co-occurrence patterns identified by the factor analysis are interpreted functionally as underlying dimensions of variation.
9. Dimension scores for each text are computed by summing the major linguistic features empirically grouped on each dimension; the mean dimension scores for each register are then compared to analyze the salient linguistic similarities and differences among spoken and written registers.
10. The functional interpretation of each dimension is revised based on the distribution of spoken and written registers along the dimension.

2.3. *Choice of linguistic features included in multi-dimensional analyses*

Although the co-occurrence patterns underlying dimensions are determined empirically, those patterns depend on the prior choice of linguistic features to be used in the analysis. Most multi-dimensional analyses to date have focussed on lexical, grammatical, and syntactic features, with the goal of being as inclusive as possible. That is, any linguistic characteristic that can be interpreted as having functional associations is a candidate for inclusion in multi-dimensional analyses. Previous analyses have included:

- lexical features, such as type–token ratio and word length;
- semantic features relating to lexical classes, such as hedges, emphatics, speech act verbs, mental verbs;
- grammatical feature classes, such as nouns, prepositional phrases, attributive and predicative adjectives, past tense verbs, perfect aspect verbs, personal pronouns; and
- syntactic features, such as relative clauses, adverbial clauses, *that* complement clauses, passive postnominal participial clauses.

One characteristic of multi-dimensional analyses is that they can be extended by investigating the role of additional features in relation to previously determined dimensions. For example, Biber (1992b) analyzes the distribution and function of linguistic features marking reference and cohesion within texts, showing how these features relate to the previously identified multi-dimensional structure of English. While some cohesion features function as part of previously identified dimensions, other cohesion features co-occur in new patterns to define additional dimensions associated with the marking of reference in discourse.

Future multi-dimensional analyses could be extended to include linguistic features from additional domains, such as the frequency of various rhetorical devices or the frequency of different organizational patterns. Any text characteristic that is encoded in language and can be reliably

identified and counted is a potential candidate for inclusion. Multi-dimensional analyses to date have focused primarily on a wide range of lexical and grammatical characteristics, but these analyses could be usefully extended to include consideration of language characteristics from other linguistic levels.

2.4. *Summary of the 1988 multi-dimensional analysis of register variation in English*

As noted in the introduction, it is important to distinguish between the multi-dimensional approach to register variation and multi-dimensional studies of particular discourse domains in particular languages. Watson focuses on the multi-dimensional analysis of English register variation presented in Biber (1988); this study provides the fullest account of multi-dimensional methodology and a synchronic analysis of the relations among adult spoken and written registers.

Five major dimensions are identified and interpreted in Biber (1988: especially chapters 6–7). Each comprises a set of co-occurring linguistic features; each defines a different configuration of similarities and differences among spoken and written registers; and each has distinct functional underpinnings. The five dimensions are interpretively labeled as follows:

1. Involved versus Informational Production
2. Narrative versus Non-narrative Concerns
3. Elaborated versus Situation-Dependent Reference
4. Overt Expression of Persuasion
5. Abstract versus Non-abstract Style

The primary communicative functions, major co-occurring features, and characteristic registers associated with each dimension are summarized in Table 1. Registers differ systematically along each of these dimensions, relating to functional considerations such as interactiveness, involvement, purpose, and production circumstances; and these functions are in turn realized by systematic co-occurrence patterns among linguistic features. The Appendix provides a more concrete illustration of how the 1988 multi-dimensional analysis can be used for comparative studies of spoken and written registers.

Two major conclusions come out of the 1988 multi-dimensional analysis of register variation in English: (1) no single dimension of variation is adequate in itself to account for the range of similarities and differences among registers—rather, multi-dimensional analyses are required; and (2) there is no absolute difference between spoken and written language—

Table 1. *Summary of functions, linguistic features, and characteristic registers for the five major English dimensions identified in Biber (1988)*

Functions	Linguistic features	Characteristic registers
Dimension 1 'Involved versus informational production'		
Involved (Inter)personal focus Interactive Personal stance On-line production	1st and 2nd person pronouns, questions, reductions, stance verbs, hedges, emphatics, adverbial subordination	Conversations, personal letters, public conversations
Informational Careful production Faceless	nouns, adjectives, prepositional phrases, long words	informational exposition, e.g., official documents, academic prose
Dimension 2 'Narrative versus non-narrative concerns'		
Narrative	past tense, perfect aspect, 3rd person pronouns, speech act (public) verbs	fiction
Non-narrative	present tense, attributive adjectives	exposition, broadcasts, professional letters, telephone conversations
Dimension 3 'Elaborated versus situation-dependent reference'		
Elaborated Situation-independent reference	WH relative clauses, pied-piping constructions, phrasal coordination	official documents, professional letters, written exposition
Situation-dependent reference On-line production	time and place adverbials	broadcasts conversations, fiction, personal letters
Dimension 4 'Overt expression of persuasion'		
Overt argumentation and persuasion	modals (prediction, necessity, possibility), suasive verbs, conditional subordination	professional letters, editorials
Not overtly argumentative	—	broadcasts, press reviews
Dimension 5 'Abstract versus non-abstract style'		
Abstract style	agentless passives, *by* passives, passive dependent clauses	technical prose, other academic prose, official documents
Non-abstract	—	conversations, fiction, personal letters, public speeches, public conversations, broadcasts

rather, particular types of speech and writing are more or less similar with respect to different dimensions.

2.5. *Multi-dimensional studies since 1988*

Although categorically ignored by Watson, there have been numerous studies adopting the multi-dimensional approach since 1988. These studies can be grouped into four major research arenas:

1. Synchronic analyses of other discourse domains, additional linguistic levels, and dialect comparisons. Several studies of author style should also be included here.
2. Diachronic analyses describing the historical evolution of registers.
3. Analyses of register variation in non-western languages, and cross-linguistic register comparisons.
4. Methodological investigations focussing on corpus design and analysis issues.

Synchronic multi-dimensional analyses have been carried out to investigate a number of particular discourse domains. For example, variation in student speaking and writing has been explored in several studies: Biber (1991) provides a multi-dimensional characterization of selected elementary school reading materials; Reppen (1994) undertakes a large-scale multi-dimensional analysis of elementary school speaking and writing, identifying the dimensions of variation operative in this domain and comparing student registers within this multi-dimensional space; Conrad (1994) compares the multi-dimensional characterizations of professional academic journals and textbooks, to provide a framework for analyzing patterns of writing development among college students. Other specialized discourse domains that have been described using multi-dimensional techniques include Connor-Linton's (1989) analysis of Soviet–American cross-cultural communication; White's (1994) analysis of variation in the language of job interviews; and Biber and Finegan's (1994c) description of the patterns of variation across sub-sections within medical research articles (viz., Introduction, Methodology, Results, and Discussion). Multi-dimensional analysis has also been used by Biber (1987) to compare written registers in British and American English.

Two specialized linguistic domains have been explored using multi-dimensional analyses: discourse complexity features (Biber, 1992a) and linguistic features marking reference and cohesion within texts (Biber, 1992b). Finally, three previous studies of author style (in addition to Watson's study) have been undertaken using the multi-dimensional framework: Connor-Linton's (1988) comparison of authors writing on

nuclear issues; Opas' (1990) analysis of Samuel Beckett's writing; and Biber and Finegan's (1994b) comparison of the written styles of Swift, Defoe, Addison, and Johnson across registers.

The multi-dimensional approach has also been used to study diachronic patterns of register variation in English. For example, Biber and Finegan (1989) traces the development of fiction, essays, and letters from 1650 to the present with respect to three linguistic dimensions; this study interprets the observed patterns of change relative to the changing purposes and readership of written texts over this period. Biber and Finegan (1992) adds an analysis of dialogue in drama and dialogue in fiction. Atkinson (1992) studies the development of five sub-registers of medical academic prose from 1735 to 1985, while Atkinson (1993) analyzes the evolution of scientific research writing, as represented in the *Philosophical Transactions of the Royal Society of London* from 1675 to 1975.

Third, the multi-dimensional approach has been used to investigate the patterns of register variation in non-western languages. Three such languages have been studied to date: Besnier's (1988) analysis of Nukulaelae Tuvaluan; Kim's (1990; Kim and Biber, 1994) analysis of Korean; and Biber and Hared's (1992a, 1992b, 1994) analysis of Somali. Taken together, these studies provide the first comprehensive investigations of register variation in non-western languages. Biber (1995) synthesizes these studies, together with the earlier multi-dimensional analyses of English, to explore cross-linguistic patterns of register variation, and to raise the possibility of cross-linguistic universals governing the patterns of discourse variation across registers.

Finally, there have been a number of publications that focus directly on methodological considerations relating to the multi-dimensional approach (for example, Biber, 1990, 1993a, 1993b, 1994; and Biber and Finegan, 1991). These studies present theoretical discussions supported by numerous empirical investigations, addressing issues such as:

- The design of representative text corpora: How many text samples are required? How long should text samples be? What sampling procedures should be used to select texts in a corpus? Are composite texts acceptable? What range of register diversity is required? What kinds of linguistic analyses are possible given different corpus designs?
- The differing definitional bases of registers/genres and text types (cf. Biber, 1989) and the validity of both kinds of text distinctions: To what extent are registers and text types well-defined in linguistic terms? In situational terms? How stable are quantitative linguistic analyses across text samples from different registers?
- The linguistic features to be included in a multi-dimensional analysis and the nature of form/function correspondences: What linguistic

features should be included? How are linguistic features identified? At what level should feature classes be defined? What functional associations do individual features have?

- The validity of the textual dimensions identified in a given language: To what extent are multi-dimensional analyses replicated across text corpora of differing sizes? To what extent are multi-dimensional analyses replicated across text corpora that are restricted in the range of register diversity? What kinds of evidence are considered in the functional interpretation of a dimension? To what extent are these functional interpretations supported by confirmatory factor analysis (cf. Biber, 1992a).

In sum, the multi-dimensional approach has been subjected to rigorous testing through direct methodological investigations as well as applications in a wide range of discourse domains and languages. Most of these studies have been published in refereed journals and books with international distributions, so that they are readily available for anyone to examine. These studies have repeatedly confirmed the reliability of the methods and the validity of the major findings, and they have consistently resulted in insights that would not have been possible without a similar synthesis of analytical techniques.

However, while ignoring all multi-dimensional studies since 1988 (except for Opas [1990]), Watson claims to have 'discovered' several 'inadequacies' and 'questionable' aspects of multi-dimensional analyses. I turn now to a discussion of the methodological criticisms raised in Watson's paper.

3. The methodological adequacy of the multi-dimensional approach

In earlier papers (e.g., Biber, 1990, 1993b), I note that much discussion of methodological issues relating to corpus linguistics, although often framed in fairly dogmatic terms, is based on intuitive judgements rather than empirical investigations. I regard this practice as unfortunate, since the primary strength of a corpus-based approach is that it enables direct empirical tests of our intuitive hunches—and it often shows that our prior expectations are not correct.

Watson, however, falls into this same trap. In his methodological critiques, Watson uses the language of empirical science (e.g., 'I have discovered ...'; 'I find ...'; 'I have shown ...') with none of the substance. That is, nearly all of his strongly-stated assertions about methodology are based on personal impressions rather than empirical investigation.

Regrettably, many of those judgements reflect fundamental misunderstandings of the multi-dimensional approach and of research techniques in computational/corpus linguistics generally.

In the following sections, I first consider issues relating specifically to the 1988 multi-dimensional analysis of register variation in English, focussing on the functional interpretation of textual dimensions (section 3.1). Then in section 3.2, I address issues relating to the automatic and interactive grammatical tagging of texts, and the role that such analyses play in the multi-dimensional approach. Finally in section 3.3 I briefly take up issues relating to corpus design.

3.1. *The statistical identification and functional interpretation of textual dimensions*

One of the major 'inadequacies' cited by Watson is the 'assumed connection between the communicative function of particular linguistic features and [the] qualitatively determined dimensions of style' (1994: 280). In other places, Watson refers to the 1988 multi-dimensional analysis as comprising 'seven perceived dimensions' (p. 243), and 'what [Biber] perceives to be the underlying dimensions of speech and writing' (pp. 242–243). Watson's impressions concerning the methodological procedure used for the identification and interpretation of dimensions can be summed up in the following quotations:

> Biber has made assumptions with regard to the underlying functional dimensions of speech and writing in English. He...has tried to substantiate his interpretations with statistical co-occurrence... (p. 245)

> ...his [Biber's] original qualitative assumptions regarding the classification of dimensions and their respective communicative functions are presented as verified, where in fact they are not. (p. 282)

These statements underscore Watson's failure to understand two basic aspects of multi-dimensional methodology:

First, the dimensions are determined quantitatively, not qualitatively; thus they represent groupings of linguistic features that have been identified empirically as co-occurring in texts across registers.

Second, the procedural order of analysis is to first identify the dimensions statistically, using factor analysis to determine the groupings of linguistic features that frequently co-occur. Subsequently, these groupings are interpreted in functional terms, to assess the underlying communicative forces associated with each cluster of linguistic features.

Watson misunderstands both of these points. The dimensions in the 1988 multi-dimensional analysis are not determined on perceptual or

functional grounds; they are rather the groupings of linguistic features that in fact frequently co-occur across the texts analyzed. Related to this, it is not the case that the statistical techniques are used in an attempt to substantiate previously determined dimensions. Rather, the statistical techniques actually determine the groupings of linguistic features that define each dimension. As a second step, these groupings are interpreted in functional terms. This latter step is explicitly interpretive, based on earlier functional studies carried out by dozens of scholars (cited in Biber [1988]) together with analysis of the co-occurring features in texts from multiple registers.[3]

In addition, Watson raises several criticisms of the functional interpretations themselves, specifically in relation to Dimensions 2, 4, and 5. These criticisms are claimed to reflect general shortcomings in the 1988 multi-dimensional analysis:

Dimension 2:

...I have discovered several other inadequacies of a more specific nature. I find dimension 2: narrative versus non-narrative to be insufficient in its description of just what narrative entails. It fails to account for differing forms and methods of narration and representations of speech and thought commonly found within prose literature. Biber's model, along this dimension, seems to assume uniformity in a text... (p. 282)

Dimension 4:

In addition, I have argued that Biber seems to have incorrectly labelled dimension 4 as overt expression of persuasion....As a result of this oversight Biber has made too narrow a selection of features... (p. 282)

Dimension 5:

Dimension 5, abstract versus nonabstractness, also raises certain methodological questions regarding Biber's model. This dimension is meant to account for abstractness, yet the dimension does not quantify the number of abstract nouns.... In addition, this dimension does not measure personification, metaphor or imagery. (pp. 282–283)

Rather than representing problems with the 1988 multi-dimensional analysis, the above statements reflect further basic misunderstandings on Watson's part. These misunderstandings relate to: (1) the analytical procedures used to identify dimensions, (2) the empirical bases for the functional interpretations, and (3) assumptions concerning the 'uniformity in a text'.

First, Watson again shows here that he does not understand the analytical procedure used to identify dimensions. It is not the case that dimen-

sions are assigned functions in any a priori manner. Further, it is not the case that linguistic features are assigned to dimensions by the researcher because they supposedly serve some pre-determined function. Specifically, Dimension 2 is not defined a priori as representing narration; Dimension 4 is not defined a priori as representing persuasion; Dimension 5 is not defined a priori as representing abstractness—I did not pre-select the features to be included on any of these dimensions. Rather, each of these dimensions represents groupings of linguistic features that in fact co-occur frequently in texts. The defining linguistic features of each dimension are first identified through statistical analysis of these quantitative co-occurrence relations; subsequently each grouping of features is interpreted in functional terms.

The fact that a dimension fails to discriminate among particular subregisters of English does not represent a shortcoming of the analysis. Each of the dimensions is a powerful predictor of register differences (see Biber, 1988: 126–129), but at the same time, each dimension fails to discriminate among some registers—when the linguistic characterizations of the registers along a dimension are in fact quite similar. Thus, one of the major findings of multi-dimensional analyses is that assessments of register similarities and differences must be made in a multi-dimensional space, because a pair of registers can be quite similar along one dimension and quite different along another.

For example, telephone conversations and academic prose are quite similar in their Dimension 2 characteristics (see Table 1): they share a preference for present tense verbs combined with a relative absence of past tense verbs, perfect aspect verbs, third person pronouns, and the other negative features grouped on this dimension. However, in terms of Dimension 1 characteristics (see Table 1), these registers represent opposite extremes, with conversations making extremely frequent use of linguistic features associated with on-line, involved production, and academic prose making extremely frequent use of linguistic features associated with informational, careful production. Similar comparisons could be made for other registers along other dimensions.

Second, Watson fails to understand the empirical basis for the functional interpretations given in Biber (1988). Analyzing the functional basis of each dimension is an overtly interpretive step in the multi-dimensional methodology, and as such it is open to extension and refinement. However, such revisions require an appropriate empirical foundation.

The goal of the functional interpretations is to assess the communicative functions and situational characteristics that are shared by the set of co-occurring linguistic features on a dimension and to then assign the

dimension an interpretive label. The functional interpretation of each dimension is based on three converging sources of information: (a) Consideration of previous micro-analyses of individual linguistic features carried out by dozens of scholars. The findings from these earlier studies are compared to assess the functions that are likely to be shared by the co-occurring features comprising the dimension. (b) Comparison of the relations among registers along the dimension, to assess the kinds of communicative and situational distinctions being made by the grouping of linguistic features. (c) Text analyses of the grouping of linguistic features comprising the dimension, considering the communicative purposes and circumstances associated with the features in multiple texts from a wide range of registers.

Researchers must be careful to avoid reifying functional interpretations. Interpretations always require additional testing and confirmation in further analyses. For example, Biber (1992a) uses covariance structure analysis and confirmatory factor analytic techniques to test the relative strengths of several multi-dimensional models of discourse complexity hypothesized on the basis of the 1988 multi-dimensional analysis.

It is also the case that multi-dimensional analyses can be extended to include additional linguistic features (see section 2.3 above). For example, Biber (1992b) shows how a number of referential and cohesion features relate to the previously identified multi-dimensional structure of English. However, the absence of such extensions does not invalidate previous multi-dimensional analyses, and no additional study conducted to date disconfirms the basic findings concerning the patterns of register variation documented in Biber (1988).

Most importantly, revisions of the functional interpretations must be based on empirical analyses comparable to the original 1988 study. That is, multi-dimensional interpretations cannot be tested or extended based solely on micro-analysis of a few individual texts from a single register, as in the case of Watson's study. Rather, since dimensions represent the linguistic features that co-occur across the full range of spoken and written registers, they must be interpreted in those terms—considering the role that the grouping of features serves in texts from the same range of registers.

Third, it is not correct to claim that multi-dimensional analyses assume uniformity within a text. In fact, in the 1988 discussion of Dimension 2 (interpreted as relating to narrative concerns) there is an analysis of a formal speech that has an embedded personal narrative (Biber, 1988: 141–142). This text has an intermediate characterization along Dimension 2 precisely because it is not uniform; it rather shifts between narrative and non-narrative purposes and discourse styles.

Further, it has been pointed out repeatedly that texts and registers can be analyzed at any level of generalization (e.g., Biber, 1988: chapters 8 and 9; Biber, 1990; Biber, 1994; Biber and Finegan, 1991). For example, Biber and Finegan (1994c) explicitly compares the multi-dimensional characteristics of the different sections typically found within medical research articles (viz., Introduction, Methodology, Results, and Discussion sections—cf. Figure 1 discussed in the Appendix). Far from assuming uniformity within texts, multi-dimensional analyses can be used to document the systematic patterns of variation that occur within texts.

3.2. *Computational 'tagging' of linguistic features*

Many of the critical statements that Watson makes about the computational procedures in the multi-dimensional approach are untrue or inaccurate. These statements can be grouped into three major categories: (1) assertions that Watson tested the adequacy of the computational procedures; (2) assertions that the computational procedures are 'unacceptable', 'inadequate', and/or 'inaccurate'; and (3) assertions that Watson tagged texts and that his procedures are more accurate and reliable than those used in Biber (1988). These issues are discussed in turn below.

3.2.1. *Assertions that Watson tested the adequacy of the computational procedures.* Watson repeatedly makes statements that either imply or directly assert that he has demonstrated certain inadequacies with the computational tagging procedures used in Biber (1988). For example: 'I have shown that Biber's tagging procedure is unacceptable...' (p. 281).

These claims are simply not true. Watson did not 'show' anything about my tagging procedure. He never had access to any of the tagging programs that I have developed, and he gives no indication that he has ever seen a text that was analyzed by my tagging programs. It is certainly not the case that Watson tested the tagging procedure on the texts in his small corpus.

3.2.2. *Assertions that the computational procedures are 'unacceptable', 'inadequate', and/or 'inaccurate'.* Watson often states supposed inadequacies in the multi-dimensional computational procedures as if they were proven truths; for example:

Overall judgements:
- 'Biber's tagging procedure is unacceptable...' (p. 281)
- 'Biber's programme is not without serious faults.' (p. 251)
- 'Biber's inadequate tagging procedure' (p. 240)

Assessments of the computer-based dictionary:
- 'the computerized base dictionary is inadequate...' (p. 281)
- 'His computerized dictionary is far from complete...' (p. 251)

Specific judgements concerning the accuracy of the grammatical tagger:
- '...the very high incidence of mistakes in tagging.' (p. 281)
- '...the tagging procedure often fails to correctly recognize linguistic features.' (p. 251)

Longer quotations:

> ...besides simple lexical nonrecognition or incorrect recognition, the majority of Biber's mistakes occur because the computer is unable to correctly assess the contextual boundaries, both syntactic and semantic, if the normal clause structure has been even moderately altered. (p. 253)

> In fact, it seems that there are two types of error within Biber's approach to tagging texts. Firstly, there are those errors which occur because there may be a 'bug' in the system, in that the algorithmic formulas are inadequate.... Secondly, a more serious error may exist in Biber's apparent belief that each word belongs to a discrete category whereas the category that a word belongs to in a sentence depends upon how that word is used in the sentence, it does not depend upon its abstract category, for example 'solid'. In other words Biber's basic approach leads to contextual errors. (p. 252)

Given the total absence of any empirical investigations testing the reliability or accuracy of the tagger, it is amazing that Watson imagined himself qualified to make such repeated, detailed assertions of inadequacy and error. The situation is worse in that previous studies have shown that these assertions are simply not accurate.

In support of these sweeping generalizations, Watson refers only to Opas (1990). In the late 1980's, Opas approached me to ask if I would be willing to tag a corpus of texts written by Samuel Beckett. I agreed to do this, and Opas used these analyses as the basis for her (1990) dissertation analyzing the multi-dimensional characteristics of Beckett's prose style.

Because much of Beckett's prose is experimental in nature, it was necessary to be especially careful in the interactive editing of tagged texts. Opas (1990: 14–15) summarizes the kinds of problems found in her tagged texts as follows:

> It seems that while the program [i.e., the 1988 grammatical tagger] may be useful for analyzing a wide range of text types, many modern literary texts can cause problems, since they often have unusual sentence structures; indeed, *some of the sentences are difficult even for the ordinary reader to analyze*. It is interesting to note that *the items which are particularly difficult to analyze are in texts that do*

not have 'normal' clause structure, but which are 'stripped' of those syntactic markers that usually help people to understand a message. [Emphases added]

Opas then goes on to make the following recommendations (Opas, 1990: 14–15):

This is not to say that one cannot use this form of analysis on literary texts; I merely wish to make the point that in using the analysis, one must keep these types of problems in mind since they can skew the results if they are not dealt with and they can significantly increase the time needed to apply the techniques.

Opas's conclusions are important cautions based on first-hand experience. It is certainly the case that any computer-based tagger will make mistakes when processing sentences that 'are difficult even for the ordinary reader to analyze'! However, these cautions provide no basis for the claims of inadequacy presented by Watson.

Watson's assessments of inadequacy are inaccurate or misleading in three respects:

First, the fact that the tagger used for the 1988 analysis produced some mis-tags does not mean that the subsequent multi-dimensional analysis was flawed in any respect. All automatic grammatical taggers of English make mistakes, because English has many lexical and grammatical ambiguities. To correct those mis-tags, all texts must be edited interactively before running subsequent multi-dimensional analyses. In fact, automatic computational tools are not required for multi-dimensional analysis at all; given sufficient time and energy, it would be possible to complete a multi-dimensional analysis of the texts in a one-million-word corpus entirely by hand. However, the use of a computer-based tagger greatly facilitates this process.

The current version of the tagger automatically identifies 231 different grammatical distinctions (represented by different tags or tag combinations). Spot-checks of the tagger show that it correctly assigns tags over 97 per cent of the time in fiction texts and over 95 per cent of the time in academic prose texts. The most problematic type of analysis is distinguishing among the various grammatical functions of *-ing* participles: as gerund, adjective, complement clause, adverbial clause, or postnominal modifying clause. (*-ing* participles in main clause progressive aspect constructions are not problematic.) *-ed* participles are also occasionally problematic due to the large number of grammatical distinctions they can represent: simple past tense verb, passive verb, perfect aspect verb, adverbial clause, postnominal modifying clause, or adjective. The tagger also flags occurrences of words that do not occur in the on-line dictionary when the morphology and grammatical context cannot be used to assign a grammatical tag with a high degree of certainty.

Following the automatic tagging, interactive programs are used to target those features that are most likely to be mis-tagged, such as *-ing* and *-ed* participles when they are not clearly functioning as main clause progressive aspect, perfect aspect, or passive voice verbs. Flagged words that do not occur in the on-line dictionary can also be checked interactively. It is thus possible to achieve very high accuracy in the final tagged texts (over 99 per cent) by checking only the small set of linguistic features that are most likely to be mis-tagged.

Second, Watson is not correct in claiming that the 1988 tagger was 'inadequate', 'unacceptable', etc. In fact, that version of the tagger was highly accurate, achieving rates well over 90 per cent across texts from a wide range of registers. It included analysis of a broader range of linguistic characteristics than most other taggers developed at the time, and it was robust across a wider range of text categories than most other taggers. (For example, the 1988 tagger was not dependent on orthographic conventions for punctuation, so that it could process both spoken and written texts.) The 1988 tagger has been used to analyze registers ranging from conversation to dense legal prose; historical texts ranging from 1650 to the present; adult texts and children's texts; native language and EFL texts.

Similarly Watson is not correct in claiming that the on-line dictionary used in the 1988 tagger is 'inadequate'. In fact, the dictionary included 50,624 lexical entries. Given that English contains hundreds of thousands of words, Watson is correct in asserting that the dictionary is 'far from complete'. But given that the dictionary comprised the most common words in English, that it was one of the largest lexical databases used in any tagger at the time, and that it was supplemented by morphological analyzers to determine the grammatical category of unknown words, the dictionary was more than 'adequate' for these purposes.

All taggers can be improved, and in the early 1990s I completed a three-year programming effort to adapt the 1988 tagger to run on DOS-based desktop computers. In the process, the tagger was redesigned so that it now includes both rule-based and probabilistic components, allowing identification of an even wider range of linguistic characteristics, and achieving an even more robust handling of different kinds of language (with accuracy rates typically well over 95 per cent). Of course, both taggers make errors. But most of those errors can be corrected by checking a relatively few linguistic characteristics using interactive editing tools. Overall, both taggers are highly accurate in their grammatical analyses, greatly facilitating empirical text studies.

Third, Watson here shows yet another fundamental misunderstanding

of the multi-dimensional analytical procedures, in this case claiming that the 1988 tagger does not handle word class ambiguities:

> a more serious error may exist in Biber's apparent belief that each word belongs to a discrete category... In other words Biber's basic approach leads to contextual errors. (p. 252)

This criticism indicates that Watson simply did not read, or at least understand, the description of the tagger in Biber (1988: 211–245). Of the 50,624 lexical entries in the computer-based dictionary, 3,440 words were listed as grammatically ambiguous—for example, *account* as a verb or a noun; *abstract* as an adjective, noun, or verb. In addition, all *-ing* and *-ed* participial forms were treated as grammatical ambiguities. A major part of the computer code in the tagger was dedicated to resolving these ambiguities, based on the surrounding linguistic context.

3.2.3. *Assertions by Watson that his procedures are more accurate and reliable than those used in Biber (1988)*

Despite his repeated criticisms of the multi-dimensional approach, and his fundamental misunderstandings of the analytical procedures used in the approach, Watson attempts to carry out a multi-dimensional analysis of author style. Although he provides no details of how he actually identified linguistic features in texts, Watson does make several general claims about his own procedures for analyzing grammatical features in texts. Many of these statements are unsubstantiated and apparently incorrect; these can be grouped into three categories: (1) the assertion that he 'tagged' texts; (2) the assertion that he followed the procedures in Biber (1988); (3) the assertion that his procedures are more accurate and reliable than those used in Biber (1988). For example:

- '...my own more accurate procedure for tagging texts' (p. 240)
- 'I strictly adhered to Biber's algorithmic formulas' (p. 253)
- 'I have changed the program used to count the features but I have not changed the definition of features.' (p. 253)
- 'By changing the program I improved the reliability of the resultant "taggings".' (p. 253)
- 'I have strictly adhered to Biber's model with one exception; the linguistic features under observation have been counted with the aid of a different computer program, a more effective program less susceptible to error.' (p. 254)

First of all, Watson seemingly did not use a tagger for his analysis at all. He instead reports using a concordancing program called WordCruncher, which he mistakenly claims enables grammatical tagging. WordCruncher

is a very powerful tool for corpus-based text analysis, displaying the occurrences of a target word in context (i.e., generating a concordance), and compiling various frequency counts relating to the distribution of words. WordCruncher, and related concordancing programs such as TACT, are very 'friendly': while they require little knowledge of computers, they produce linguistic analyses that would be difficult to accomplish by hand.

However, there is no sense in which a concordancing program can be considered a tagger. Concordancing programs contain no dictionary that can be used to identify the grammatical class of words, and therefore they have no means of identifying basic grammatical categories, such as nouns, verbs, adjectives, or adverbs. While they can be used to identify words that are morphologically regular, such as adverbs that end in *-ly*, it is not possible to even begin to analyze the full set of words commonly occurring in any grammatical class. For example, a concordancing program offers no means of identifying the many common adverbs that do not end in *-ly*, such as *again*, *always*, *down*, *just*, *often*, *once*, and *very*. In addition, a concordancing program provides no way to distinguish between the many adjective/adverb ambiguities in English; for example, *back*, *close*, *far*, *fast*, *hard*, *high*, *later*, *low*, *near*, *quick*, *slow*, *sure*, *wide*. These problems are even worse for the grammatical classes that are less clearly marked morphologically: nouns, adjectives, and verbs. Finally, concordancing programs provide no mechanisms for synthesizing word class disambiguation with algorithms that identify larger syntactic structures (such as relative clauses, complement clauses, etc.).

Second, Watson incorrectly asserts that he followed the algorithms for identifying linguistic features in Biber (1988). One of the two specific examples that he provides of this is the algorithm for WH relative clauses on object positions (Biber, 1988: 235):

XXX + YYY + N + WHP + ZZZ

where N stands for any noun, WHP stands for a WH relative pronoun, XXX stands for any word except for the verbs *ask* and *tell* (to exclude indirect WH questions), and ZZZ stands for any word that is not an adverb, auxiliary verb, or main verb (to exclude relativization on subject position). Watson (pp. 250–251) incorrectly identifies the variable ZZZ in this algorithm as standing for any other word (when it actually stands for words that are *not* adverbs, auxiliaries, or verbs).

More importantly, it is simply not possible to replicate the algorithms in Biber (1988) using a concordancing program. The two specific algorithms cited by Watson both illustrate this point. The algorithm for WH relative clauses depends on an accurate identification of all nouns,

adverbs, and verbs—in absolute terms and then matched in the correct sequence.

The other algorithm cited by Watson identifies perfect aspect verbs:

A) declarative structures:
HAVE+(ADV)+(ADV)+VBN
B) interrogative structures:
HAVE+N/PRO+VBN

Part A of this algorithm, for perfect aspect verbs in declarative clauses, identifies sequences of any form of the verb *have*, followed optionally by one or two adverbs, followed by a past participle verb. Part B of the algorithm, for perfect aspect verbs in questions, identifies sequences of any form of the verb *have*, followed by any noun or pronoun, followed by a past participle verb.

This is a relatively simple algorithm, and a concordancing program could be used to identify parts of it: most forms of the verb *have* (apart from *'d* contractions, which are grammatically ambiguous), adverbs that end in *-ly*, the various subject pronouns, and regular past participles that end in *-ed*. However, a concordancing program could not be used to disambiguate among *-ed* forms functioning as adjectives, past tense verbs, and past participles. Further, it could not identify nouns, adverbs which are not marked morphologically (see discussion above), and irregular past participles which are less clearly marked morphologically (a class of c. 250 verbs including many of the most common verbs in English, such as *taken*, *sent*, *cut*). Thus a large proportion of perfect aspect verb constructions would be overlooked by a concordancing program.

Similar problems would be encountered with almost every one of the algorithms used in the 1988 tagging program. These problems are even more serious for complex grammatical constructions, such as *that* complement clauses headed by verbs or adjectives, with the *that* present and omitted. For example;

I hope [*that*] *you understand.*
I'm sure [*that*] *he went.*

Concordancing programs provide no mechanisms for distinguishing among the various uses of the word *that* (for example, as demonstrative, demonstrative pronoun, relative pronoun, and complementizer); and identifying complement clause constructions where the complementizer *that* has been ommitted is even more difficult.

Third, Watson provides no details of how he actually attempts to identify linguistic features in texts, apart from noting that he used

Wordcruncher 'to count the frequency of the linguistic features' (p. 252), together with his assertions that his procedures exactly follow the algorithms in Biber (1988) and that they are more accurate and reliable than those used in Biber (1988).

The claim for greater accuracy is based on yet another fundamental misunderstanding of the methodological procedures used in Biber (1988). Watson discusses the occurrence of errors resulting from the tagging program as if these errors were tolerated in the final multi-dimensional analysis. This is not the case; the final counts used in all multi-dimensional analyses must be accurate. Use of a tagging program enables automatic generation of accurate counts in c. 90–95 per cent of all cases, but interactive editing is done to insure accurate counts in the remaining 5–10 per cent of cases where errors might occur (see, for example, Biber, 1988: 216–217).

In addition, given Watson's heavy reliance on a concordancing program, this claim of greater accuracy is not likely to be correct: the algorithms in the 1988 tagger cannot be replicated by a concordancing program, due to the absence of an on-line dictionary and strict limitations on the complexity of logical operations possible in concordancing programs. For this reason, it is likely that Watson's counts were considerably less complete and less accurate than those produced by any reasonable tagging program, including the one used in Biber (1988).

3.3. *Corpus design issues*

The final area where Watson claims to have identified basic inadequacies with the 1988 multi-dimensional analysis of English concerns the corpus design; for example:

> Yet the more serious fault lies with the corpus.... The use of these corpora raise certain unanswered questions. Are these corpora indicative of written and spoken English as a whole?...Is it correct to only use British English as a base corpora for comparison...? (p. 281)

> Biber, having a less than adequate corpus should have been more tentative in his claims. (p. 282)

> In fact the LOB corpus is too small...and as a result interpretations made from this corpus may be skewed. At the very best they will only represent written British English for that very year. (p. 282)

Here, as elsewhere, Watson simply asserts his opinions as if they had been substantiated by empirical evidence. In fact, Watson undertook no

empirical investigation of corpus design issues at all—he did no analyses based on the LOB, Brown, and London-Lund Corpora used in earlier multi-dimensional studies; and he completed no empirical studies relating to corpus design issues in any corpus.

However, several studies of this type have been reported elsewhere. For example, the statistical investigations reported in Biber (1990, 1993a, 1993b) show that size and register-diversity are equally important considerations in corpus design. These studies further show that earlier corpora, such as the LOB, Brown, and London-Lund (which together comprise 2.5 million words of text), are more than adequate for many quantitative linguistic analyses, including reliable distributional analyses of common linguistic features and the linguistic characteristics of genres and text types. In particular, these studies have shown that the dimensions of variation identified in the 1988 multi-dimensional analysis can be replicated in much smaller corpora, if the corpus represents the same range of register variation.

Biber (1987) further shows that there are only small differences in the multi-dimensional characteristics of written registers in British English and American English. Far from being valid for only a single point in time, Biber and Finegan (1989, 1992) and Atkinson (1992) show that the dimensions identified in the 1988 multi-dimensional analysis can be usefully applied to describe the patterns of register variation in English over the last three centuries. Finally, Biber (1995) shows that a comparison of multi-dimensional analyses across languages and cultures uncovers striking similarities in both the underlying dimensions of variation and the relations among registers, suggesting the operation of underlying form/function associations tied to basic aspects of human communication.

4. Summary and conclusion

Watson, while apparently being unaware of all multi-dimensional research studies published since 1988 (except for Opas [1990]), feels qualified to assert critical opinions as if they were established research findings. The above sections have shown that he adopts this same attitude in his claims about all major methodological aspects of multi-dimensional analyses: the interpretive techniques, statistical techniques, computational techniques, and corpus design considerations. Unfortunately Watson is wrong in two regards: his opinions do not represent findings, and in fact, most of his criticisms have been shown to be unfounded by empirical investigation.

Appendix. Illustrative multi-dimensional comparison of ten spoken and written registers using two dimensions from Biber (1988)

To illustrate the multi-dimensional characterization of registers more concretely, Figure 1 presents the differences among ten spoken and written registers within the two-dimensional space defined by Dimension 1 (Involved versus Informational Production) and Dimension 5 (Abstract versus Non-abstract Style). The defining linguistic features and functional interpretations for these dimensions are summarized in Table 1 (in section 2.4).

The register characterizations on Figure 1 reflect different relative frequencies of the co-occurring linguistic features comprising each dimension. For example, medical research articles and scientific prose have the largest negative scores on Dimension 1 (scores between −18 and −25 on the vertical axis); these scores represent very frequent occurrences of nouns, attributive adjectives, long words, and prepositional phrases, together with markedly infrequent occurrences of private verbs, *that*-deletions, contractions, etc. (i.e., the 'involved' features on Dimension 1). Medical and scientific prose also have the largest positive scores on Dimension 5 (scores between 5 and 9 on the horizontal axis); these scores reflect very frequent occurrences of conjuncts, agentless passives, past participial adverbial clauses, and *by*-passives.

At the other extreme, conversations have the largest positive score on Dimension 1, reflecting very frequent occurrence of the features relating to 'involvement' and 'on-line production', such as first and second person pronouns, questions, contractions, stance verbs, hedges, and emphatics. At the same time, this score for conversations along Dimension 1 reflects the relative absence of 'informational' features such as nouns, attributive adjectives, and long words. Conversations also have the largest negative score on Dimension 5, reflecting the near complete absence of features such as conjuncts, agentless passives, and *by*-passives.

As can be seen from Figure 1, these ten registers are strikingly different in their linguistic characteristics, even within this two-dimensional space. For example, scientific prose and medical prose are quite different from the other eight registers in being extremely informational (Dimension 1) and abstract/passive in style (Dimension 5). Figure 1 further shows that there are systematic but much smaller differences among the four subregisters within medical research articles; for example, Methods sections are marked as the most informational and abstract type of prose considered here.

At the other extreme, conversation is marked as being extremely involved and non-passive/non-abstract in its linguistic characteristics.

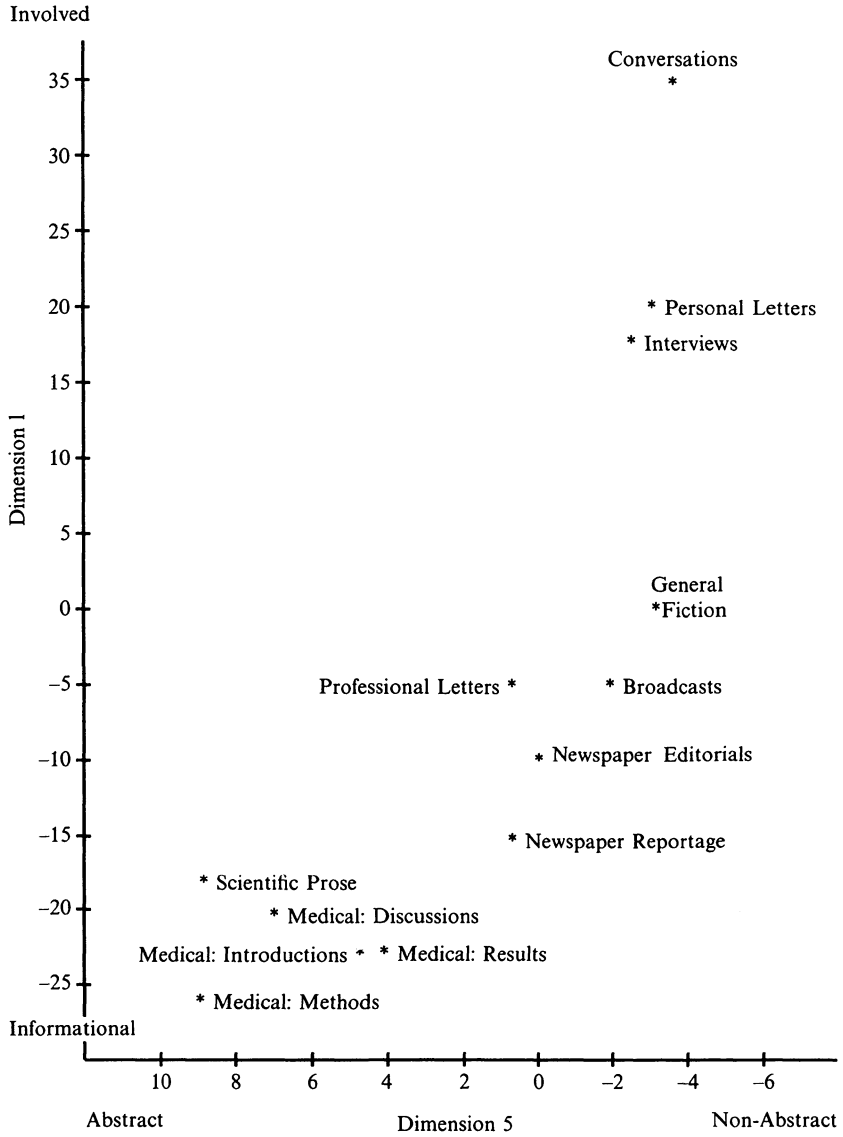

Figure 1. *Linguistic characterization of ten spoken and written registers with respect to Dimension 1: 'Involved versus Informational Production', and Dimension 5: 'Abstract versus Non-Abstract Style'*

Table 2 *F scores and correlations for dimension score differences across 23 spoken and written registers (df = 22,459)*

Dimension	F value	Probability	r^2
1	111.9	$p < 0.001$	84.3%
2	32.3	$p < 0.001$	60.8%
3	31.9	$p < 0.001$	60.5%
4	4.2	$p < 0.001$	16.9%
5	28.8	$p < 0.001$	58.0%

Personal letters and interviews have similar characteristics, but the characterizations are not nearly as extreme as in conversations.

When all five major dimensions are considered, the differences among registers are even more notable. Table 2 shows that each of the dimensions are significant and important predictors of linguistic variation among registers. It is important to emphasize here that the register categories were not considered when the dimensions were originally identified; rather, the dimensions represent the linguistic co-occurrence patterns that actually occur across texts, regardless of their register category. However, as both Figure 1 and Table 2 show, there are important differences across registers with respect to each dimension.

Notes

* I would like to thank Susan Conrad, Edward Finegan, Randi Reppen, and Teun van Dijk for their many helpful comments on earlier drafts of this paper.

1. In earlier work (e.g., Biber, 1985, 1986, 1988), I have referred to this analytical approach as the multi-feature/multi-dimensional approach. Watson repeatedly refers to the approach as 'multifeatured', although this is not a term that I have used myself.
2. At one level these two goals are diametrically opposed, in that the multi-dimensional analysis of Nyoongah's style follows an extended critique of the 1988 multi-dimensional analysis of English. The following concluding statements from Watson's article provide a striking illustration of this basic contradiction:

> To summarise, Biber's multifeatured/multidimensional model is not without its faults. His choice of corpora and preparation of corpora is questionable. As a result of this I believe comparative studies with Biber's corpus are dubious... Even so, the interpretations along the dimensions are based on generalizations and lack precise microscopic units of quantification.
>
> However, as a direct result of this inadequacy Biber's model offers the possibility of measuring broad changes within texts or across texts for the six dimensions that have been examined. This lends itself very well to the investigation of style shift, both synchronically and, in particular, diachronically. (1994: 283)

Watson cannot have it both ways—the 1988 multi-dimensional analysis of English

cannot be appropriately used for insightful comparative analyses if it is fundamentally flawed.

3. At some points Watson even appears to misunderstand the fact that multi-dimensional analyses are inherently *multi*-dimensional; for example, he incorrectly claims that the 1988 multi-dimensional analysis of English supports 'the hypothesis that there is a single spoken/written continuum' (1994: 240).

References

Aijmer, K. (1984). *Sort of* and *kind of* in English conversation. *Studia Linguistica* 38: 118–128.

Altenberg, B. (1984). Causal linking in spoken and written English. *Studia Linguistica* 38: 20–69.

Atkinson, D. (1992). The evolution of medical research writing from 1735 to 1985: The case of the Edinburgh Medical Journal. *Applied Linguistics* 13: 337–374.

— (1993). A historical discourse analysis of scientific research writing from 1675 to 1975: The case of the *Philosophical Transactions of the Royal Society of London*. Unpublished Ph.D. Dissertation, University of Southern California.

Beaman, K. (1984). Coordination and subordination revisited: Syntactic complexity in spoken and written narrative discourse. In *Coherence in Spoken and Written Discourse*, D. Tannen (ed.), 45–80. Norwood, N.J.: Ablex.

Besnier, Niko (1988). The linguistic relationships of spoken and written Nukulaelae registers. *Language* 64: 707–736.

Biber, Douglas (1985). Investigating macroscopic textual variation through multi-feature/multi-dimensional analyses. *Linguistics* 23: 337–360.

— (1986). Spoken and written textual dimensions in English: Resolving the contradictory findings. *Language* 62: 384–414.

— (1987). A textual comparison of British and American writing. *American Speech* 62: 99–119.

— (1988). *Variation across Speech and Writing*. Cambridge: Cambridge University Press.

— (1989). A typology of English texts. *Linguistics* 27: 3–43.

— (1990). Methodological issues regarding corpus-based analyses of linguistic variation. *Literary and Linguistic Computing* 5: 257–269.

— (1991). Oral and literate characteristics of selected primary school reading materials. *Text* 11: 73–96.

— (1992a). On the complexity of discourse complexity: A multidimensional analysis. *Discourse Processes* 15: 133–163.

— (1992b). Using computer-based text corpora to analyze the referential strategies of spoken and written texts. In *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82, Stockholm, 4–8 August 1991*, Jan Svartvik (ed.), 213–252. Berlin: Mouton.

— (1993a). Representativeness in corpus design. *Literary and Linguistic Computing* 8: 243–257.

— (1993b). Using register-diversified corpora for general language studies. *Computational Linguistics* 19: 219–241.

— (1994). An analytical framework for register studies. In D. Biber and E. Finegan (eds.), 31–56.

— (1995). *Dimensions of register variation: A cross-linguistic comparison*. Cambridge: Cambridge University Press.

Biber, D. and Finegan, E. (1989). Drift and the evolution of English style: A history of three genres. *Language* 65: 487–517.

— (1991). On the exploitation of computerized corpora in variation studies. In *English Corpus Linguistics: Studies in Honour of Jan Svartvik*, K. Aijmer and B. Altenberg (eds.), 204–220. London: Longman.

— (1992). The linguistic evolution of five written and speech-based English genres from the 17th to the 20th centuries. In *History of Englishes: New Methods and Interpretations in Historical Linguistics*, M. Rissanen et al. (eds.), 688–704. Berlin: Mouton.

— (eds.) (1994a). *Sociolinguistic Perspectives on Register*. New York: Oxford University Press.

— (1994b). Multi-dimensional analyses of authors' styles: Some case studies from the eighteenth century. *Research in Humanities Computing 3*, D. Ross and D. Brink (eds.), 3–17. Oxford: Oxford University Press.

— (1994c). Intra-textual variation within medical research articles. In *Corpus-based Research into Language*, N. Oostdijk and P. de Haan (eds.), 201–222. Amsterdam: Rodopi.

Biber, D. and Hared, M. (1992a). Dimensions of register variation in Somali. *Language Variation and Change* 4: 41–75.

— (1992b). Literacy in Somali: Linguistic consequences. *Annual Review of Applied Linguistics* 12: 260–282.

— (1994). Linguistic correlates of the transition to literacy in Somali: Language adaptation in six press registers. In D. Biber and E. Finegan (eds.), 182–216.

Chafe, W. L. (1982). Integration and involvement in speaking, writing, and oral literature. In *Spoken and Written Language: Exploring Orality and Literacy*, D. Tannen (ed.), 35–54. Norwood, N.J.: Ablex.

Coates, Jennifer (1983). *The Semantics of the Modal Auxiliaries*. London: Croom Helm.

Connor-Linton, Jeff (1988). Author's style and world-view in nuclear discourse: A quantitative analysis. *Multilingua* 7: 95–132.

— (1989). Crosstalk: A multi-feature analysis of Soviet-American spacebridges. Unpublished Ph.D. dissertation, University of Southern California.

Conrad, Susan (1994). Variation in academic writing: Textbooks and research articles across disciplines. Paper presented at the annual conference of the American Association of Applied Linguistics, Baltimore.

Kim, Yong-Jin (1990). Register variation in Korean: A corpus-based study. Ph.D. dissertation, Department of Linguistics, University of Southern California.

Kim, Y-J. and Biber, D. (1994). A corpus-based analysis of register variation in Korean. In D. Biber and E. Finegan (eds.), 157–181.

Opas, L. L. (1990). Aspects of style in Samuel Beckett's literature. Unpublished Ph.D. dissertation, Oxford University.

Reppen, Randi (1994). Variation in elementary student writing. Ph.D. dissertation, Department of English, Northern Arizona University.

Schiffrin, D. (1981). Tense variation in narrative. *Language* 57: 45–62.

— (1987). *Discourse markers*. Cambridge: Cambridge University Press.

Tannen, D. (1982). Oral and literate strategies in spoken and written narratives. *Language* 58: 1–21.

Thompson, S. A. (1983). Grammar and discourse: The English detached participial clause. In *Discourse Perspectives on Syntax*, F. Klein-Andreu (ed.), 43–65. New York: Academic Press.

Tottie, G. (1986). The importance of being adverbial: Focusing and contingency adverbials in spoken and written English. In *English in Speech and Writing: A Symposium*, G. Tottie and I. Bäcklund (eds.), 93–118. (Studia Anglistica Upsaliensia 60.) Stockholm: Almqvist and Wiksell.

Watson, G. J. (1994). A multidimensional analysis of style in Mudrooroo Nyoongah's prose works. *Text* 14: 239–285.

White, Margie (1994). Language in job interviews: Differences relating to success and socioeconomic variables. Ph.D. dissertation, Department of English, Northern Arizona University.

Douglas Biber is an associate professor in the Applied Linguistics Program (English Department) at Northern Arizona University. His research interests focus on linguistic variation (lexical, grammatical, and discourse) across registers, both in English and cross-linguistically.