
Speaking and Writing in the University: A Multidimensional Comparison

DOUGLAS BIBER

*Northern Arizona University
Flagstaff, Arizona, United States*

SUSAN CONRAD

*Portland State University
Portland, Oregon, United States*

RANDI REPPEN

*Northern Arizona University
Flagstaff, Arizona, United States*

PAT BYRD

*Georgia State University
Atlanta, Georgia, United States*

MARIE HELT

*California State University
Sacramento, California, United States*

The dozens of studies on academic discourse carried out over the past 20 years have mostly focused on written academic prose (usually the technical research article in science or medicine) or on academic lectures. Other registers that may be more important for students adjusting to university life, such as textbooks, have received surprisingly little attention, and spoken registers such as study groups or on-campus service encounters have been virtually ignored. To explain more fully the nature of the tasks that incoming international students encounter, this article undertakes a comprehensive linguistic description of the range of spoken and written registers at U.S. universities. Specifically, the article describes a multidimensional analysis of register variation in the TOEFL 2000 Spoken and Written Academic Language Corpus. The analysis shows that spoken registers are fundamentally different from written ones in university contexts, regardless of purpose. Some of the register characterizations are particularly surprising. For example,

classroom teaching was similar to the conversational registers in many respects, and departmental brochures and Web pages were as informationally dense as textbooks. The article discusses the implications of these findings for pedagogy and future research.

Teachers, textbook authors, and test developers are constantly faced with decisions regarding the language forms, topics, and functions to include in ESOL materials. Unfortunately, few empirical linguistic descriptions are available to inform these decisions. As a result, language professionals rely on intuitions and anecdotal evidence of how speakers and writers use language. Despite the value of intuitions in materials development, intuitions about language use often turn out to be wrong (see Biber & Conrad, 2001; Biber & Reppen, in press). Comprehensive linguistic descriptions are not all that materials developers need, but such descriptions provide essential information for making principled pedagogical decisions. The research reported in this article contributes empirically based linguistic description intended to inform materials development for university-level English language instruction.

TESOL professionals are aware of the special demands of academic reading and writing, especially in relation to textbooks, research papers, and student essays and term papers. Teachers also recognize the importance of academic listening skills, which are required for success in the classroom. However, considerably less attention has been directed toward other university registers,¹ such as study groups, office hours, and course packs. Institutional registers that may be particularly important for students to negotiate include written registers—such as handbooks, catalogues, program Web pages, and course syllabi—and spoken registers, such as service encounters with the registrar or departmental staff and the classroom management talk of instructors at the beginning of class sessions. Little is known about the linguistic characteristics of these registers, so it is not surprising that most programs and textbooks do not address the language skills required to handle them.

¹The term *register* here is a cover term for any language variety defined in situational terms, including the speaker's purpose in communication, the topic, the relationship between speaker and hearer, spoken or written mode, and the production circumstances (see Biber, 1994, 1995; Conrad & Biber, 2001). Registers can be described at any level of generality (Biber, 1994). For example, *methodology sections in chemistry research articles* is a highly specified register; *academic prose* is a very general register (unspecified for many characteristics).

Because registers are defined in situational rather than linguistic terms, texts from the same register can have extensive linguistic differences. Some registers, like official documents, are very consistent in their linguistic characteristics; texts from other registers, like fiction, can be very different in their linguistic characteristics. An alternative approach is to define text categories in linguistic terms, called *text types* in previous multidimensional studies (see Biber, 1995, chapter 9).

To better understand the nature of the language that incoming international (and domestic) students encounter in the university, and ultimately to help students develop the language skills required, TESOL professionals need a comprehensive linguistic description of all spoken and written registers used at the university. This article reports results of the most comprehensive linguistic analysis of academic language to date. The study draws on quantitative linguistic analysis of the TOEFL 2000 Spoken and Written Academic Language (T2K-SWAL) Corpus, which was designed to represent the full range of spoken and written registers used at U.S. universities (e.g., classroom teaching, office hours, study groups, textbooks) as well as in the major academic disciplines (e.g., humanities, natural sciences) and academic levels (lower division, upper division, and graduate).

BACKGROUND

Approaches to Academic Discourse Analysis

The many studies on academic discourse published over the past 20 years have been undertaken from a variety of perspectives (see, e.g., the extensive survey of research in Grabe & Kaplan, 1996). Many of these studies adopt a rhetorical or social/historical perspective, describing the rhetorical structure of academic texts and the way the practices of researchers in particular discourse communities shape the conventions of academic genres. Most studies focus on written scientific or medical prose (see, e.g., the book-length studies by Atkinson, 1999; Bazerman, 1988; Berkenkotter & Huckin, 1995; Gilbert & Mulkay, 1984; Halliday & Martin, 1993; Swales, 1990; Valle, 1999).

Other studies describe the surface linguistic characteristics of academic texts, again focusing mostly on written academic registers, especially academic research articles in science or medicine. The hedging devices used in academic texts have been particularly well researched (see, e.g., Crompton, 1997; Grabe & Kaplan, 1997; Holmes, 1988; Hyland, 1994, 1996a, 1996b). Several other studies document the special classes of verbs used in research articles (e.g., Hunston, 1995; Thompson & Ye, 1991; Williams, 1996) and the complex noun phrase structures typical of scientific prose (e.g., Halliday, 1988; Love, 1993; Varantola, 1984). Other analysts have described specialized linguistic features, such as imperatives (Swales et al., 1998), conditionals (Ferguson, 2000), personal pronouns (Kuo, 1999), existential *there* (Huckin & Pesante, 1988), politeness markers (Myers, 1989), citation patterns (Salager-Meyer, 1999), procedural vocabulary (Marco, 1999), and collocational frames (Marco, 2000). At the other extreme, as part of a corpus-based

reference grammar, Biber, Johansson, Leech, Conrad, and Finegan (1999) describe the grammatical features in academic prose in comparison with those in conversation, fiction, and newspaper reportage. Atkinson (1992, 1996, 1999) and Conrad (1996, 2001) describe the characteristics of professional written registers with respect to a large number of co-occurring linguistic features (see the section Multidimensional Analysis below).

Few studies have described the linguistic characteristics of spoken academic discourse. The numerous studies of the rhetorical organization of classroom discourse (see, e.g., Cazden, 1988) have focused for the most part on discourse markers and other relatively fixed lexical chunks (e.g., Chaudron & Richards, 1986; Flowerdew & Tauroza, 1995; Khuwaileh, 1999; Nattinger & DeCarrico, 1992; Strodt-Lopez, 1991) or on the overall discourse organization of the lecture (see, e.g., the papers in Flowerdew, 1994). Carson, Chase, Gibson, and Hargrove (1992) discuss how undergraduates are required to integrate written and spoken registers, specifically by reading textbooks to prepare to listen to lectures. Even fewer studies have described the linguistic characteristics of other spoken registers common in university life. A recent exception to this generalization is Cutting's (1999) analysis of the conversations of a group of postgraduate students.

This review of past studies in academic discourse reveals a focus on written academic prose and academic lectures, with the overwhelming majority of the research on the technical research article (in science or medicine). Past work has neglected other registers important for students, such as textbooks and spoken registers (e.g., study groups or on-campus service encounters).

To help international university students develop the language skills they need, TESOL professionals might benefit from a comprehensive linguistic description of all university spoken and written registers, including textbooks and classroom teaching experiences. Equally important, although perhaps less obvious, are the "gatekeeping" registers, like university catalogues, departmental Web pages, course syllabi, class management talk (in which instructors describe course requirements and expectations), and service encounters (in which newly arrived students interact with office staff to accomplish the business of becoming a student). In sum, the TESOL profession needs fuller linguistic descriptions as the basis for ESL and English for academic purposes (EAP) materials that represent the full extent of ESL students' future university tasks.

Multidimensional Analysis

Previous research on academic discourse has been limited in part simply because researchers have been interested in specific registers or linguistic features rather than the overall patterns of register variation. In addition, more comprehensive investigations have not been feasible until recently. The combined use of computer programs for automated language processing and representative text corpora enables such comprehensive investigations (cf. Biber, Conrad, & Reppen, 1998). Corpus-based analysis allows for the following essentials:

1. the adequate representation of naturally occurring discourse. Corpora can include representative text samples from a variety of academic registers, allowing for analyses based on long passages from each text, multiple texts from each register, and a full range of spoken and written registers.
2. (semi-)automatic linguistic processing of texts using computational processing. This allows comprehensive linguistic characterization of a text through description of a wide range of linguistic features.
3. reliable, accurate quantitative analyses of linguistic features. Because computers do not become bored or tired, they count a linguistic feature in the same way every time it is encountered.
4. the possibility of cumulative results and accountability. Subsequent studies can be based on the same corpus of texts, or additional corpora can be analyzed using the same computational techniques.

Taking advantage of these potentials for linguistic analysis, our study of academic registers used a quantitative, corpus-based technique called multidimensional (MD) analysis. MD analysis was developed to discover and interpret the patterns of linguistic variation found in a corpus of texts. Early researchers in sociolinguistics (e.g., Ervin-Tripp, 1972) argued that linguistic features work together in texts as constellations of co-occurring features (rather than as individual features) to distinguish among registers. Although this theoretical perspective is widely accepted, before the availability of corpus-based techniques, linguists lacked the methodological tools required to analyze these co-occurring features. MD analysis uses multivariate statistical techniques to investigate the quantitative distribution of linguistic features across texts and text varieties and to analyze linguistic co-occurrence by identifying underlying dimensions of variation through a statistical factor analysis.

The dimensions identified in MD analysis have both linguistic and functional interpretations. The linguistic content is a group of features (e.g., nouns, attributive adjectives, prepositional phrases) that co-occur

with a markedly high frequency in texts. On the assumption that co-occurrence reflects shared functions, analysts interpret the co-occurrence patterns to assess the situational, social, and cognitive functions most widely shared by the linguistic features. For example, the frequent co-occurrence of first-person pronouns, second-person pronouns, hedges, and emphatics in conversational texts is interpreted as reflecting directly interactive situations and a primary focus on personal stance and involvement (see below).

Biber (1988) identified five main dimensions of variation in a general corpus of spoken and written registers. He used factor analysis to identify the groups of linguistic features associated with each dimension (i.e., the linguistic features that co-occur in texts with markedly high frequencies; see Table 1). The dimensions represent the co-occurrence distributions of 67 linguistic features across 481 spoken and written texts of contemporary British English. The texts, taken from the Lancaster-Oslo-Bergen Corpus and the London-Lund Corpus, represent 23 major register categories (e.g., academic prose, press reportage, fiction, letters, conversations, interviews, radio broadcasts, public speeches). The factor loadings for a linguistic feature (Table 1, third column) can range from -1.0 to $+1.0$; the farther from 0.0 a loading is, the stronger the association between the feature and the dimension. Features with higher loadings are thus better representatives of the dimension underlying a factor.

Most of the dimensions consist of two groupings of features, one with positive and the other with negative loadings. The positive and negative sets represent features that occur in a complementary pattern. That is, when the features in one group occur together frequently in a text, the features in the other group are markedly less frequent in that text, and vice versa.

Interpretations of the dimensions should consider likely reasons for the complementary distribution of these two groups of features as well as the reasons for the co-occurrence pattern within each group. For example, on Dimension 1, the interpretation of the features having negative loadings is relatively straightforward because the features are relatively few in number. Nouns, word length, prepositions, type/token ratio, and attributive adjectives all have negative loadings larger than .45, and no feature has a larger loading on another factor. High frequencies of all these features indicate an informational focus and a careful integration of information in a text. These features are associated with texts that have an informational purpose and provide ample opportunity for careful integration of information and precise lexical choice.

The set of features with positive loadings on Dimension 1 is more complex, although all of these features have been associated with an involved, noninformational focus related to a primarily interactive or affective purpose and on-line production circumstances. For example,

TABLE 1
Summary of Biber's (1988) Factor Analysis

Feature	Example	Factor loading
Dimension 1: Involved versus informational production		
Positive features (involved production)		
Private verbs	think, know, believe	0.96
<i>that</i> -deletions	I think [0] he went	0.91
Contractions	can't, she's	0.90
Present tense verbs	is, likes, wants	0.86
Second-person pronouns	you	0.86
<i>do</i> as pro-verb	so <i>did</i> Sandra	0.82
Analytic negation	that's <i>not</i> likely	0.78
Demonstrative pronouns	this shows...	0.76
General emphatics	really, a lot	0.74
First-person pronouns	I, we	0.74
Pronoun <i>it</i>	I didn't like <i>it</i>	0.71
<i>be</i> as main verb	that <i>was</i> sad	0.71
Causative subordination	because . . .	0.66
Discourse particles	well, anyway	0.66
Indefinite pronouns	nothing, someone	0.62
General hedges	kind of, something like	0.58
Amplifiers	absolutely, extremely	0.56
Sentence relatives	Bob didn't study at all, <i>which is usual for him</i>	0.55
<i>wh</i> - questions	Why did you go?	0.52
Possibility modals	can, could, may, might	0.50
Nonphrasal coordination	Sally was biking last weekend <i>and</i> then she . . .	0.48
<i>wh</i> - clauses	Jill asked <i>what happened</i>	0.47
Final prepositions	the candidate that I was thinking <i>of</i>	0.43
Negative features (informational production)		
Nouns	community, case	-0.80
Word length	—	-0.58
Prepositions	of, in, for	-0.54
Type/token ratio	—	-0.54
Attributive adjectives	good, possible	-0.47
Dimension 2: Narrative versus nonnarrative discourse		
Positive features (narrative discourse) ^a		
Past tense verbs	considered, described	0.90
Third-person pronouns	he, she, they	0.73
Perfect aspect verbs	had been, has shown	0.48
Public verbs	said, explain	0.43
Synthetic negation	no answer is good enough	0.40
Present participial clauses	Having established the direction, we can now . . .	0.39
Dimension 3: Situation-dependent versus elaborated reference ^b		
Positive features (situation-dependent reference)		
Time adverbials	early, instantly, soon	0.60
Place adverbials	above, beside, outdoors	0.49
Adverbs	always, significantly	0.46

Continued on next page

TABLE 1, continued
Summary of Biber's (1988) Factor Analysis

Feature	Example	Factor loading
Negative features (elaborated reference)		
<i>wh</i> -relative clauses in object positions	something <i>which everybody can do</i>	-0.63
Pied piping constructions	the way <i>in which this happens</i>	-0.61
<i>wh</i> -relative clauses in subject positions	those <i>who retain inhibitions</i>	-0.45
Phrasal coordination	salt and pepper	-0.36
Nominalizations	extension, proposition	-0.36
Dimension 4: Overt expression of persuasion		
Positive features (overt expression of persuasion)		
Infinitives	hope to go	0.76
Prediction modals	will, would, shall	0.54
Suasive verbs	command, insist, propose	0.49
Conditional subordination	if you want	0.47
Necessity modals	must, should, have to	0.46
Split auxiliaries	should <i>really</i> be	0.44
(Possibility modals)	can, could, might	(0.37) ^c
Dimension 5: Nonimpersonal versus impersonal style ^{b, d}		
Negative features (impersonal style)		
Conjuncts	however, therefore	-0.48
Agentless passives	The same mechanism <i>was analyzed</i> on each.	-0.43
Past participial adverbial clauses	Directed by Twilling, the production is delightful.	-0.42
<i>by</i> passives	He <i>was surrounded by a ring of men.</i>	-0.41
Past participial postnominal clauses	the course <i>chosen by the large majority</i>	-0.40
Other adverbial subordinators	since, while, whereas	-0.39

Note. The table includes only features with loadings larger than ± 0.35 ; features with smaller loadings have not demonstrated strong evidence for their occurrence on the dimension.

^aNo negative features. ^bPolarity reversed; see Footnote 2. ^cFeature was not used in the computation of dimension scores. ^dNo positive features.

first- and second-person pronouns, *wh*- questions, emphatics, amplifiers, and sentence relatives can all be interpreted as reflecting interpersonal interaction and the involved expression of personal feelings and concerns. Other features with positive loadings on Dimension 1 mark a reduced surface form, a generalized or uncertain presentation of information, and a generally fragmented production of text; these include *that*-deletions, contractions, pro-verb *do*, the pronominal forms, and final (stranded) prepositions. In these cases, a reduction in surface form also results in a more generalized, less explicit content.

Overall, based on both positive and negative co-occurring linguistic features, Dimension 1 seems to represent a dimension marking affective, interactional, and generalized content (the features with positive loadings) versus high informational density and precise informational content (the features with negative loadings). Two separate communicative parameters seem to be represented here: the primary purpose of the writer/speaker (involved vs. informational) and the production circumstances (those dictated by real-time constraints vs. those enabling careful editing possibilities). Reflecting both of these parameters, the interpretive label *involved versus informational production* seems appropriate for this dimension.

The complementary groupings of features on the other factors shown in Table 1 reflect other functional relations. The interpretive labels for the dimensions (*involved versus informational production*, *narrative versus nonnarrative discourse*, *situation-dependent versus elaborated reference*, *overt expression of persuasion*, and *nonimpersonal versus impersonal style*) express the communicative function(s) they represent (see Table 2).² Biber (1988, chapters 6–7; 1995, chapters 5–7) and Conrad and Biber (2001, chapter 2) provide justification for these interpretations based on the shared communicative functions of the co-occurring linguistic features on each dimension plus the distribution of registers along each dimension.

Having defined these dimensions empirically through quantities of linguistic characteristics, we can analyze any text by computing its *dimension score*: a summation of the frequencies for those features having salient loadings on a dimension. Registers and subregisters can then be compared in terms of their mean dimension scores. Considering all five dimensions together enables multidimensional analyses of the linguistic characteristics of particular registers and the linguistic differences among registers.

Biber (1988) used these dimensions to compare and contrast a wide variety of spoken and written registers (including conversation, personal letters, fiction, and academic prose). Subsequent studies have used the dimensions to analyze academic registers in greater detail. For example, Conrad (1996, 2001) compared the multidimensional characteristics of research articles and textbooks in the academic disciplines of ecology and American history. This study provides a baseline for the study of writing development, comparing the characteristics of term papers written by students at various levels to the characteristics of professional written texts. Carkin (2001) focused on introductory textbooks and

²To facilitate comparisons across dimensions in this analysis, we reverse the polarity of Dimensions 3 and 5 as given by Biber (1988). Dimension 5 has only negative features, reflecting an *impersonal* style. Because the opposing end of this dimension has no features at all, we refer to it as *nonimpersonal* style, which is not necessarily the same as a *personal* style.

TABLE 2
Communicative Functions Represented by Biber's (1988) Dimensions

Dimension	Functions	Example
Involved versus informational production	Interactive discourse with high involvement and a focus on personal stance versus carefully produced written discourse with an informational purpose	Conversation versus academic prose
Narrative versus nonnarrative discourse	Stereotypically narrative discourse	Fiction versus conversation and academic prose
Situation-dependent versus elaborated reference	Situated reference versus elaborated, context-independent reference	Sports broadcasts versus official documents
Overt expression of persuasion	Persuasive or argumentative discourse	Editorial versus normal conversation
Nonimpersonal versus impersonal style	Focus on events and circumstances rather than the participants	Conversation and fiction versus scientific academic prose

lectures, using the dimensions for a four-way comparison of lower-division textbooks and lectures in economics and biology. Biber and Finegan (1994) compared the multidimensional profiles of the introduction-methods-results-discussion sections in medical research articles. Csomay (2000) used a modified MD analysis to investigate the characteristics of academic lectures. Other studies have used the five dimensions to track historical patterns of change within academic written registers, focusing especially on medical prose and science prose (Atkinson, 1992, 1996, 1999; Biber, 1995, chapter 8; Biber & Finegan, 1997; see also the papers in Conrad & Biber, 2001). Taken together, these studies demonstrate the power of the Biber's (1988) multidimensional framework for building descriptions of academic registers.

While continuing to investigate academic discourse, the present study takes a broader perspective than these previous investigations did. Rather than focus on a few stereotypically academic registers, we analyze a full range of registers encountered by students in university life. Some of these registers—such as classroom teaching, office hours, and study groups—are influenced by competing functional forces, for example, the need to convey information efficiently versus the restrictions of real-time (spoken) production and the need for social interaction. But how are these and other, competing functional influences reflected in the language of the texts in each register? To investigate this question, we

locate each academic register along the five register dimensions described above.

METHOD

Corpus Design and Data Collection

We designed the T2K-SWAL Corpus to be relatively large (2.7 million words) and to represent the academic registers that U.S. university students must listen to or read (see Table 3).³ The register categories chosen for the corpus reflect the spoken and written activities associated with academic life, including class sessions, office hours, study groups, on-campus service encounters, textbooks, course packs, and other campus writing (e.g., university catalogues, brochures). The sampling weight

TABLE 3
Composition of the TOEFL 2000 Spoken and Written Academic Language Corpus

Register	Texts	Words
Spoken		
Class sessions	176	1,248,811
Classroom management	40 ^a	39,255
Labs/in-class groups	17	88,234
Office hours	11	50,412
Study groups	25	141,140
Service encounters	22	97,664
Total	251	1,665,516
Written		
Textbooks	87	760,619
Course packs	27	107,173
Course management	21	52,410
Other campus writing	37	151,450
Total	172	1,071,652
Overall total	423	2,737,168

^aClassroom management texts were extracted from class session texts, so they are not included in the total text counts.

³The corpus is being used (a) for a series of linguistic investigations and (b) to provide a baseline for test materials. Related to the first purpose, we are investigating various linguistic characteristics of academic texts, including vocabulary distributions, the use of collocations and lexical bundles, grammatical characteristics, syntactic complexity, informational density, and the expression of stance. In all cases, the design of the corpus allows research to be undertaken from the perspective of register comparison. That is, each register can be studied in relation to the other academic spoken and written registers. Related to the second purpose, the corpus is being used to ensure that test stimuli represent the same range of linguistic (lexical and grammatical) complexity that students encounter regularly in academic life.

given to each register category reflects our assessment of its relative availability and importance.

To gather data, we identified and captured naturally occurring discourse at four academic sites (California State University, Sacramento; Georgia State University; Iowa State University; and Northern Arizona University). Taken together, the sites represent four U.S. regions—West Coast (California State), Rocky Mountain West (Northern Arizona), Midwest (Iowa State), and Deep South (Georgia State)—and four types of academic institutions—a teacher’s college (California State), a mid-size regional university (Northern Arizona), an urban research university (Georgia State), and a rural research university (Iowa State). Although we did not achieve full demographic/institutional representativeness, we aimed to avoid obvious skewing for these factors.

For the spoken corpus, our participants were primarily students, whom we recruited to record their academic conversations faculty. We also recruited faculty to record office hours and university staff to record service encounters. Student participants carried audiocassette recorders to capture academic speech as it occurred in the class sessions and study groups that they were involved in over a 2-week period, keeping a log of speech events and participants to the extent that it was practical. Faculty simply left cassette recorders running during their office hours (with student consent). This approach overcame the tendency for the somewhat artificial discourse that is often created by the presence of research assistants in spoken settings. We obtained high-quality, natural interactions; the main disadvantage was that we did not observe the interactions firsthand and thus could not obtain detailed information about the setting and participants.

Service encounters were recorded wherever students regularly interacted with staff to conduct the business of the university. These settings included the university bookstore, copy shop, and coffee shop; the front desk in the dormitory; academic department offices; the library information desk; the media center; and student business services.

For class sessions and textbooks, we sampled spoken and written texts from six major disciplines (business, education, engineering, humanities, natural science, and social science) and three levels of education (lower-division undergraduate, upper-division undergraduate, and graduate). Table 4 shows the breakdown of texts by discipline and level for class sessions and for textbooks. Recognizing the existence of systematic variation within each of these high-level disciplines, we also targeted specific subdisciplines (e.g., chemistry, philosophy, psychology); although these distinctions will allow for register comparisons at a more specific level in future research, we restricted the study described here to the main categories. Finally, the corpus includes various teaching styles, as

TABLE 4
Class Session Texts and Textbooks in the Corpus by Discipline and Level

Discipline and level	Class session texts		Textbooks	
	Texts	Words	Texts	Words
Business				
Lower division	8	44,418	4	29,744
Upper division	20	126,026	4	28,399
Graduate	8	66,010	7	58,078
Total	36	236,454	15	116,221
Education				
Lower division	4	26,237	2	18,601
Upper division	4	25,871	2	15,830
Graduate	8	85,135	2	15,685
Total	16	137,243	6	50,116
Engineering				
Lower division	8	45,864	3	18,629
Upper division	14	72,165	3	24,902
Graduate	8	53,156	3	28,482
Total	30	171,185	9	72,013
Humanities				
Lower division	10	65,984	6	56,324
Upper division	12	91,732	6	52,870
Graduate	9	90,946	6	54,938
Total	31	248,622	18	164,132
Natural science				
Lower division	9	48,616	6	53,564
Upper division	7	40,447	6	42,555
Graduate	9	71,810	6	48,995
Total	25	160,873	18	145,114
Social science				
Lower division	15	124,435	7	75,324
Upper division	15	107,283	7	71,182
Graduate	8	62,712	7	66,517
Total	38	294,430	21	213,023

measured by the extent of interactivity in classroom teaching, but this study did not consider such distinctions.

Course packs collected for the corpus included written texts of several types: lecture notes, study guides, detailed descriptions of assignments or experimental procedures written by the instructor, and photocopies of published journal articles and book chapters. Course management texts are mostly syllabi, but this category also includes some written assignments or exams. Finally, the category *other campus writing* included the miscellaneous written texts that students encounter on campus, such as

informational brochures about academic programs; university catalogues; Web pages describing academic programs; and informational brochures on student services, health or safety issues on campus, scholarships, and other topics. Although not often considered academic discourse, written material of this type is among the first that a prospective student receives from a university. It is ubiquitous on campus and is required reading for a prospective student attempting to navigate the maze of university requirements and services.

Data Coding

All texts in the corpus were coded with a header to identify content area and register. Spoken texts were transcribed using a consistent convention (see Edwards & Lampert, 1993), and to the extent possible speakers were distinguished and some demographic information for each (e.g., status as instructor or student) supplied in the header.

After editing all texts to ensure accuracy in transcribing and scanning, we grammatically annotated the texts using an automatic grammatical tagger (developed and revised over a 10-year period by Biber). The grammatical tags were then edited using an interactive grammar checker to ensure a high degree of accuracy for the final annotated corpus (see Biber et al., 1998, Methodology Boxes 4 and 5). For example, following is the tagged equivalent of the sentence *The dissolved components that precipitate to form these rocks are decomposed from pre-existing rocks and minerals.*

The ^ati++++
dissolved ^jj+atrb++xvbn+
components ^nns++++
that ^tht+rel+subj++
precipitate ^vb++++
to ^to++++
form ^vbi++++
these ^dt+dem+++
rocks ^nns++++
are ^vb+ber+aux++
decomposed ^vpsv++agls+xvbnx
from ^in++++
pre-existing ^jj+atrb++xvbg+
rocks ^nns++++
and ^cc++++
minerals ^nns++++
. ^.+clp+++

Data Analysis

For the quantitative linguistic comparisons of texts and registers, we used a computer program that calculated the rate of occurrence of linguistic features in each text (e.g., the number of nouns per 1,000 words). The linguistic variables analyzed in the T2K-SWAL Corpus for the purposes of the present study were the same features used for Biber's (1988) factor analysis of general spoken and written registers (summarized above); 67 linguistic variables were analyzed (see Table 1 above). In the present study, we applied the dimensions from the 1988 factor analysis to compare university spoken and written registers. That is, we analyzed the same linguistic features in the texts of the T2K-SWAL Corpus and then calculated dimension scores for those texts.

To determine the distribution of university registers along each dimension, we compared texts and registers with respect to those dimension scores. The normalized linguistic feature counts are scores that show the rate of occurrence in texts (e.g., a noun score, an adjective score). In a similar way, dimension scores (or *factor scores*) can be computed for each text by summing the scores of the features having salient loadings on that dimension. In this study, only features with loadings greater than 0.35 on a factor were considered important enough to be used in computing dimension scores. For example, we computed the Dimension 1 score for each text by adding together the frequencies of private verbs, *that*-deletions, contractions, present tense verbs, and so on—the features with positive loadings on Factor 1 (from Table 1)—and then subtracting the frequencies of nouns, word length, prepositions, and so on—the features with negative loadings.

The individual linguistic variables were standardized to a mean of 0.0 and a standard deviation of 1.0 before the dimension scores were computed. This process translates the scores for all features to scales representing standard deviation units, so that all features on a factor have equivalent weights in the computation of dimension scores (see Biber, 1988, pp. 93–97).

Once a dimension score had been computed for each text, we computed the mean dimension score for each register. Plots of these mean dimension scores allow linguistic characterization of any given register, comparison of the relations between any two registers, and a fuller functional interpretation of the underlying dimension (see, e.g., Figure 1 below). In a similar way, standard statistical procedures (such as analysis of variance [ANOVA]) can be used to analyze the statistical significance of differences among the mean dimension scores.

RESULTS

To summarize the many features of the language in the corpus, we identify the texts' positions along the five dimensions described in the Background section. This analysis shows how the university registers vary. We then explore this variation further by analyzing differences among texts associated with different disciplines and levels of study.

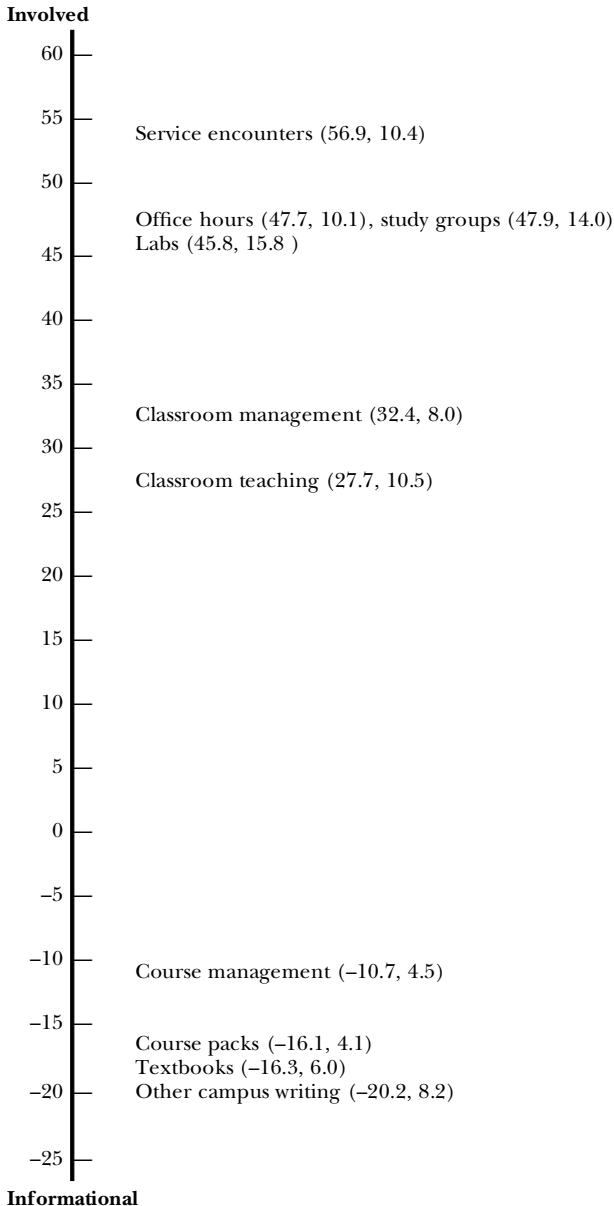
Patterns of Variation Among University Registers

To describe the variation among registers, we plotted the mean dimension scores for the 10 university registers included in the T2K-SWAL Corpus based on the combined scores of the co-occurring features in each text (see Figures 1–5 below). Registers with large positive mean scores on a particular dimension contained high frequencies of the positive features for that dimension and low frequencies of its negative features (see Table 1 above). Conversely, registers with large negative mean scores on a dimension have high frequencies of the negative features of that dimension and low frequencies of the dimension's positive features. These plots reveal several interesting findings about the linguistic characteristics of individual registers and about the patterns of variation among university registers. The statistics at the bottom of each figure report the results of an ANOVA to test for significant differences among the registers with respect to that dimension score. The r^2 value is a direct measure of strength, reporting the proportion of variance for the dimension score that can be predicted by the register distinctions. For example, Dimension 1 is a very strong predictor of register differences, with 88.9% of the variance for this dimension score predicted by register (see Figure 1). In the following subsections, we consider each dimension in turn.

University Registers Along Dimension 1: Involved Versus Informational Production

The distribution of university registers along Dimension 1 is surprising (see Figure 1). Previous multidimensional studies have interpreted Dimension 1 as a reflection of two underlying functional considerations: (inter)personal versus informational primary purpose and real-time versus careful production circumstances. Biber's (1988) study of general spoken and written registers showed considerable overlap among registers along this dimension, reflecting the complex interplay of these

FIGURE 1
Mean Scores of University Registers Along Dimension 1,
Involved Versus Informational Production



Note. $F = 401.3$; $df = 9, 453$; $r^2 = .889$; $p < .001$. The two figures in parentheses are mean scores and standard deviations, respectively.

factors. For example, prepared speeches (spoken) and fiction (written) both had Dimension 1 scores of around 0.0 (Biber, 1988, p. 128).

In contrast, spoken and written university registers were completely polarized along Dimension 1. All written registers had large negative scores, reflecting a frequent use of the negative features on Dimension 1 (e.g., nouns, long words, prepositions, attributive adjectives; see Table 1 above), coupled with the relative absence of positive features on this dimension. From a functional perspective, these negative scores indicate that the written registers are extremely informational in purpose and are produced under highly controlled and edited circumstances. Interestingly, the register of other campus writing has this same characterization even though the category is composed of nonacademic texts like brochures and university catalogues. Text Sample 1 illustrates the dense informational characteristics of textbooks, and Text Sample 2 illustrates the similar characteristics of nonacademic written materials. (Nouns are underscored, attributive adjectives are in italics, and prepositions are in uppercase letters.)

1. The formation OF a *separate* socialist bloc would insulate the East FROM the *coming economic* chaos IN the West and enhance socialist economic development. The *primary* motivation, however, was political. A *separate Eastern economic* bloc, IN the Soviet Union's view, would provide a buffer zone OF *friendly*, that is, Communist states ON its borders and would prevent Germany or other "*hostile*" Western powers FROM posing a threat OF military invasion. Furthermore, the Soviet Union would obtain access ON *favorable terms* TO the resources OF Eastern Europe—*raw materials* and capital equipment—that could be used to rebuild the Soviet Union AFTER the war and to advance its *economic* development. THROUGH wartime diplomacy, military occupation, and coups d'état, the Soviet Union established Communist satellite regimes IN all the states OF Eastern Europe. (textbook: social science [political science], upper division, tbpol2.sir)⁴
2. The Center FOR Academic Success serves students BY providing information ABOUT college programs, student professional organizations, career opportunities, campus support services, and college and University policies and procedures, including General Education advising. Referrals are made to direct students TO the most *appropriate* department when *further* information is required. Additionally, academic support is provided THROUGH study groups directed BY student tutors. Establishing *good* study habits and working WITH other students are

⁴Text samples are identified by register (type), discipline (subdiscipline), and level (if applicable). We also include the filename in the T2K-SWAL Corpus so that the larger textual context of these samples can be examined in future research.

essential FOR success IN *technical fields*. The groups are organized BY the Center FOR a variety OF engineering, computer science, math, and science courses. (other campus writing [Web page]: College of Engineering, otcatc.eng)

On the spoken end, all university registers had scores indicating high involvement, reflecting their frequent use of such features as present tense verbs, private verbs, first- and second-person pronouns, and contractions. The most surprising inclusion in this group was classroom teaching, which had a notably involved rather than informational characterization. This finding suggests that classroom teaching in U.S. universities is much more involved or interactive and less fully scripted than prepared speeches (including formal lectures) are. That is, whereas prepared speeches are carefully scripted and have a relatively informational characterization along Dimension 1, classroom teaching is more spontaneous and therefore is characterized by a greater use of features marking personal involvement and real-time production, as shown in Text Sample 3.

3. Teacher: I guess uh . . . I'm trying to think of other levels here but the question that you have to ask is what kind of resources do you have internally? And what do you have to get externally? And what are you good at and what are you not good at? To be able to really do good innovation to get products out and I contend also to have good e-commerce sites and good e-business sites is that you have some combination of some of these things. The more you have a whole set of resources it's more likely that you are going to have a competitive advantage, and then the question is which one of these do you have and which one of these are you going to find [two unclear syllables] in other ways? What's a way if you don't have these resources, what's a way to get some?

Student: What are the [unclear words] to the right?

Teacher: Oh I'm sorry - marketing, manufacturing, I'm not even sure [unclear words] quiet economy, I was just trying to think of other things on the fly - uh you may - do you guys have any things that I've missed here? (classroom teaching: business, upper division, busmgleudhg104)

Text Sample 3, with its dense use of first- and second-person pronouns, *wh-* questions, and present tense verbs, illustrates the highly involved/interactive, relatively unplanned nature of typical university teaching. Classroom teaching is informational as well as involved, as reflected in the use of nouns, adjectives, and prepositional phrases in the text (e.g., *kind of resources*; *good e-commerce sites*; *good e-business sites*; *marketing*, *manufacturing* . . . *quiet economy*). However, much of the

information is inexplicit. For example, forms like pronouns (e.g., *which one of these do you have and which one of these are you going to find; some of these things*) and *wh*-questions and clauses (e.g., *what do you have to get; what are you good at*) are commonly used instead of more precise noun phrases. When nouns are used, they are often vague in reference (e.g., *thing: I was just trying to think of other things*).

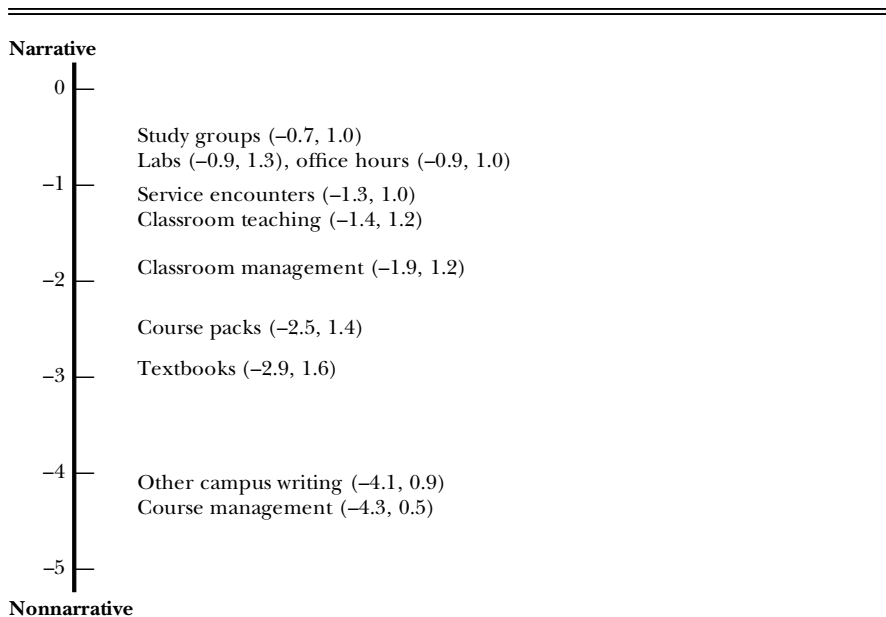
It is not surprising that explicitly interactive academic registers, like office hours and study groups, show the characteristics of involvement to an even greater extent than university teaching does. That is, although these registers have a primary informational purpose, the demands of interpersonal communication and real-time production have a much stronger influence in determining their linguistic characteristics. Text Sample 4 illustrates the highly involved, unplanned nature of a typical study group.

4. 1: You understand what he's saying? How to do it? Just take that thing
2: Not quite
1: [unclear] copy it a number of times
2: Pretty much.
1: So, what you think? that this doesn't work?
2: What? Well, I think when he does that he's got that array, that val array. I think you got an array with all that stuff in there. Like val one is equal to whatever that thing is.
1: Uh-huh, then what is tend do the one that he
2: He just then uses the array. He does the exact same thing as the array. Just -
1: Lots of copies.
2: Yeah,
1: Copies of this?
2: No. It had, it had like, did you see it?
1: No.
2: It was like a - like T I S R R, one, and I don't know - like where he got that address. He just had like, go back, said like this is equal to -
1: Oh, oh, oh, oh, oh, oh
2: I think I can. This works on fifteen, right? (study group: engineering, upper division, engcpsgudgi091)

University Registers Along Dimension 2: Narrative Versus Nonnarrative Discourse

Relative to Dimension 1, the academic registers show little variation along Dimension 2 (see Figure 2). For the most part, university registers are characterized by the absence of narrative features. The written

FIGURE 2
Mean Scores of University Registers Along Dimension 2,
Narrative Versus Nonnarrative Discourse



Note. $F = 34.7$; $df = 9, 453$; $r^2 = .408$; $p < .001$. The two figures in parentheses are mean scores and standard deviations, respectively.

registers—especially other campus writing and course management—have especially large negative scores, representing an extremely low quantity of narrative features. Text Sample 2 above illustrates the absence of these features in a departmental Web page.

The absence of narrative features in textbooks is surprising, given the widespread perception that textbook authors from many disciplines rely heavily on narratives. However, this finding agrees with earlier MD studies of disciplinary writing (especially Conrad, 2001; see also Biber et al., 1998, pp. 158–163), which have shown that even textbooks for disciplines with a focus on the past do not typically rely on narrative discourse. That is, although these textbooks include some narratives written entirely in the past, present tense discussions of implications are much more common. The narrative sections may be perceptually salient, but they do not account for much of the total discourse in university-level textbooks. (Reppen, 2001, shows, however, that elementary school social science textbooks are much more narrative in general; see also Biber et al., 1998, pp. 180–188.)

In contrast to the norms for written registers, spoken university

registers—especially study groups, office hours, and labs—show a greater tendency to use narrative features to some extent, resulting in Dimension 2 scores closer to 0.0. These scores reflect a mixing of purposes in these registers, including discussion and explanation of academic topics coupled with a recounting of past classroom teaching. Text Sample 5 illustrates this recounting in an interaction from an office hours meeting.

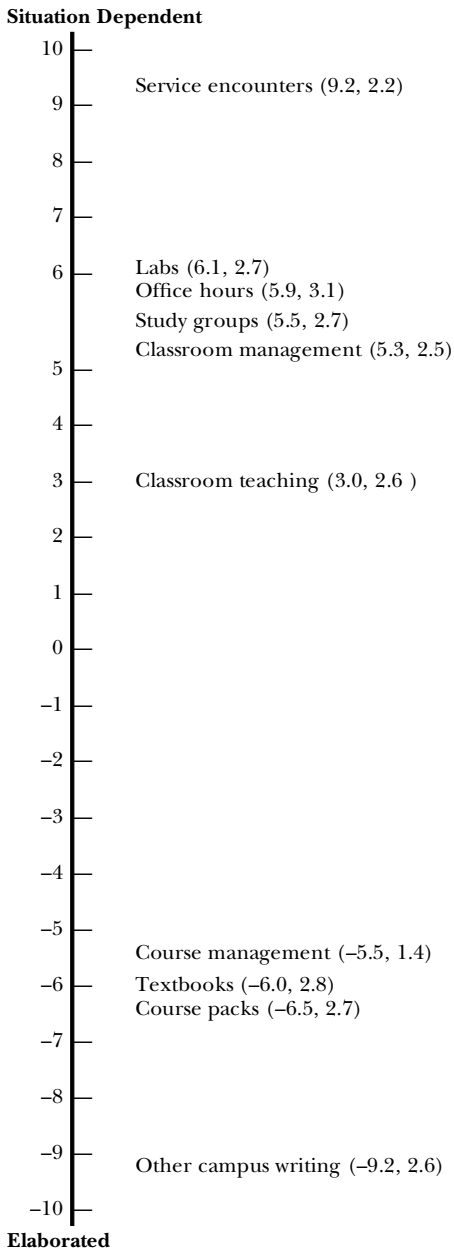
5. Teacher: What did I call the foreign policy in the twenties?
Student: That would be my next question like
Teacher: Well that's my next questio- I beat you to it. Are you an athlete?
Student: Me?
Teacher: Yeah
Student: No [laughing]
Teacher: Oh ok - what, uh, remember I said we picked and chose it was like going to the cafeteria
Student: Yeah that was, uh, picking we just picked the fights that we wanted to be in
Teacher: What did I call that? - a la carte, remember?
Student: A la carte
Teacher: Yeah, ok, so we picked and chose - well why did we pick and choose? Because we hadn't joined the the League of Nations in which all of this would be decided
Student: Ok (office hours: humanities [history], humioh_n071)

University Registers Along Dimension 3: Situation-Dependent Versus Elaborated Reference

Dimension 3, plotted in Figure 3, is similar to Dimension 1 in that it defines an absolute polar distinction between written and spoken university registers. Positive scores along this dimension represent a frequent use of time and place adverbials, reflecting *situation-dependent reference*, whereas large negative scores represent the frequent use of *wh*-relative clauses, phrasal coordination, and nominalizations, interpreted as *elaborated reference*.

Spoken university registers with large positive scores on Dimension 1 can be considered situation-dependent in some ways, as illustrated in Text Samples 3, 4, and 5 above: These texts display the dense use of pronouns (e.g., *it*, demonstrative pronouns such as *this* and *those*) and generalized nouns (e.g., *thing*) that take their meaning from the larger textual and situational context. These same registers commonly rely on directly situated reference, as reflected in their frequent use of adverbials that refer directly to the time and place of the speech event. Service encounters are especially marked for these features, but the academic interactive registers also use them frequently. Text Sample 6, from an

FIGURE 3
Mean Scores of University Registers Along Dimension 3,
Situation-Dependent Versus Elaborated Reference



Note. $F = 234.6$; $df = 9, 453$; $r^2 = .823$; $p < .001$. The two figures in parentheses are mean scores and standard deviations, respectively.

office hours session, illustrates the frequent direct references to time and place common in this register (such references are underscored).

6. Teacher: Yeah
Student: Um back right before two twenty
Teacher: Yeah
Student: Or what time?
Teacher: Yeah yeah and the class starts at two twenty
Student: Ok
[another student enters]
Teacher: Yes
Student: Um I actually - I missed the test, I fell asleep today
Teacher: You fell asleep - What are you doing at two twenty?
Student: Um taking the test hopefully
Teacher: Come into room one oh one
Student: I'll be there sir, thank you. (office hours: business, busbaoh__n156)

At the other extreme, written registers like textbooks and course packs are characterized by a dense use of relative clauses and phrasal coordination, reflecting styles of referring that are minimally dependent on the situational context. Interestingly, other campus writing is by far the most elaborated register along this dimension. Text Sample 7, from an anthropology Web page, illustrates the dense use of these features (relative pronouns are underscored; instances of phrasal coordination are in italics).

7. The Master of Arts program in Anthropology is designed for students who plan to continue their graduate studies in anthropology at the Ph.D. level *as well as* for students who plan to pursue any of the numerous opportunities for graduate anthropologists, such as in private research, foreign service, education, *and* government.

The program centers on a core of general requirements designed to provide each student with a graduate level exposure to the broad discipline of anthropology, with an emphasis on general methodology *and* the ways in which problems are conceptualized *and* approached in at least three of the interrelated subdisciplines. [. . .]

[. . .] After the graduate interview, the student forms his/her Thesis Committee, which is composed of a graduate advisor *and* at least one additional member of the Anthropology faculty. (other campus writing [department Web page]: anthropology, otcatc.ant)

It is noteworthy that other campus writing is marked as the most "literate" register along Dimensions 1 and 3, reflecting an extremely dense concentration of complex nominal constructions, such as nouns, attributive adjectives, prepositional phrases, and technical vocabulary on

Dimension 1, and *wh*-relative clause constructions on Dimension 3. This register is, in a sense, the front door to the university, as it includes the texts that all students must read to understand the procedures and requirements of university programs. It is therefore interesting that these texts should be more structurally complex than the content taught in university courses. This finding gives empirical linguistic support to the old saw about university catalogues: “If a student can read it, admit her or him. If she or he can understand it, give her or him a degree.”

University Registers Along Dimension 4: Overt Expression of Persuasion

The defining features on Dimension 4 include several modal and semimodal verbs related to prediction (e.g., *will*, *would*, *be going to*) and necessity (e.g., *must*, *should*, *have to*). In addition, this dimension includes suasive verbs (e.g., *command*, *demand*, *insist*) and conditional subordination. These co-occurring features have been interpreted as reflecting an overtly persuasive style. Registers like newspaper editorials use these features to a greater extent than other registers do, but most previous MD studies found no register to be especially marked for these co-occurring features.

In contrast, all spoken university registers use these features relatively frequently (see Figure 4), and two of these registers—classroom management and office hours—are especially marked for their dense use.⁵ In addition, written course management shows a dense presence of these features. What these registers seem to have in common is their focus on behavior modification. Simply put, they try to persuade students to perform required tasks according to course specifications. Text Samples 8 and 9 illustrate these features (modal verbs are underscored).

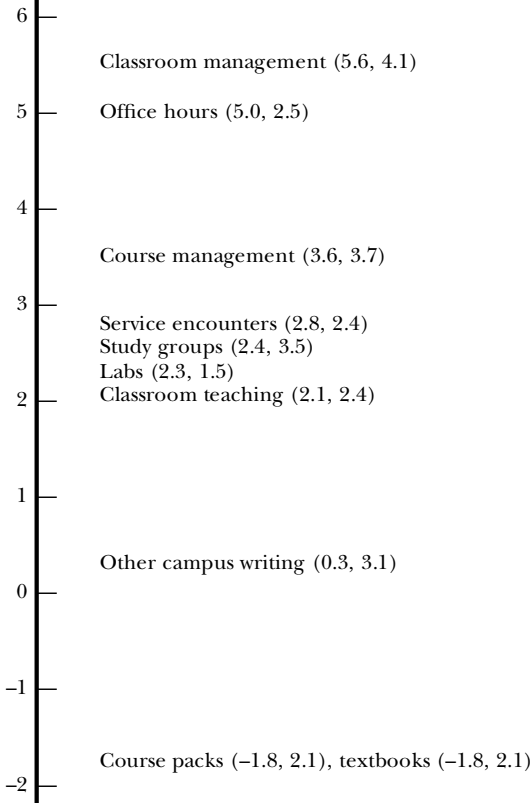
8. Teacher: OK now the presentation in here [clears throat] [four unclear syllables] will be next week Wednesday
Student: And
Teacher: And OK here's what you're gonna have to do - you're gonna have to show either using the navigation tool or use the story board type PowerPoint presentation to show how you designed and set up your system board, alright ?
Student: [whistling]

⁵ Interestingly, university service encounters are more persuasive/argumentative in their Dimension 4 characterization than general U.S. service encounters are (as described in Helt, 2001). We attribute this difference to the inclusion of extended information-seeking interactions in our corpus of university service encounters (e.g., at the library, student business services, academic departments) in contrast to the reliance on store interactions in most other corpora of service encounters.

FIGURE 4

Mean Scores of University Registers Along Dimension 4, Overt Expression of Persuasion

Overtly persuasive



Not overtly persuasive

Note. $F = 35.3$; $df = 9, 453$; $r^2 = .412$; $p < .001$. The two figures in parentheses are mean scores and standard deviations, respectively.

Teacher: You know what your (flows) are which windows are gonna come up whether you chose to use primary verses secondary dialogs and why.
 Student: Maybe that should be my PowerPoint
 Teacher: What ?
 Student: Maybe [five unclear words]
 [laughter]
 [instructor and students talking at the same time]
 Teacher: And then after you show that then you're gonna have to bring your model up and show it running right ? Somewhat that is and uh

Student A: You have to have a model

Student B: Somewhat

Teacher: Yeah

[laughter] (classroom management [in-class discussion of course assignment]: engineering, engcslegrhn217.txt)

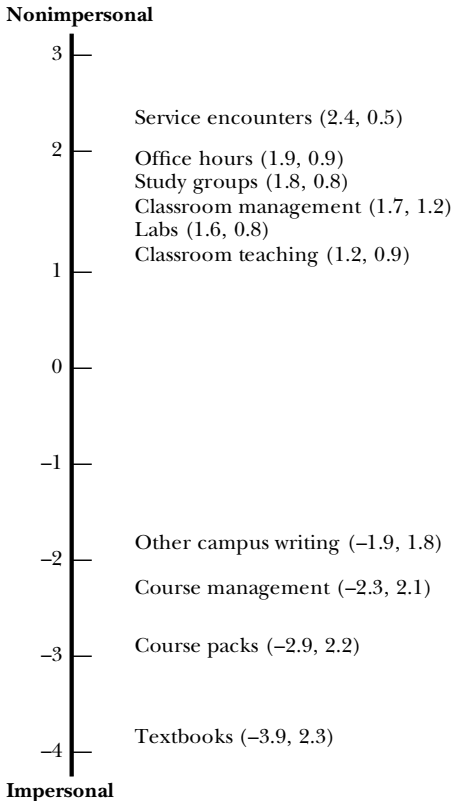
9. The outline provided is tentative but should be adequate enough to give you a reference for the order topics will be covered and a reasonable idea of the pace the materials will be covered. Students are expected to come to class prepared to actively participate in the learning process. As in any professional organization, absences should be justified and promptness standard procedure. Your homework should be done with pride and submitted on time. Late homework will not be accepted. Every person who contributes to the solution will get the same score. Only one solution is to be submitted from the group. Persons not contributing will receive no credit. (natural science, course management: meteorology, course syllabus, upper-division undergraduate, cmnsc2.syl)

University Registers Along Dimension 5: Nonimpersonal Versus Impersonal Style

Along Dimension 5, texts vary in their use of passive constructions, including main-clause verb phrases and postnominal modifiers, and in their use of certain kinds of connecting words. Similar to the patterns observed along Dimensions 1 and 3, spoken and written registers show an absolute distinction along Dimension 5: All spoken registers in the corpus are marked by the absence of these passive constructions whereas all written registers use passive features frequently (see Figure 5). These features are especially common in textbooks, as illustrated in Text Sample 10 (passive constructions are underscored; conjuncts are in italics).

10. The hypothetical spectrum of dimethyltrifluoroacetamide presented at the end of Chapter 1 may have suggested that NMR spectroscopy is employed for the detection of magnetically different nuclei in a compound. For at least two reasons this is not the case. *Firstly*, experimental considerations make such an application difficult, if not impossible, since conditions and techniques must be modified to measure the resonance frequencies of different nuclei. *Secondly*, the elemental composition of organic compounds can be determined far more easily and accurately by other techniques such as elemental analysis or mass spectrometry. The significance of NMR spectroscopy in chemistry is therefore not based on its ability to differentiate between elements, but on its ability to distinguish a particular nucleus with respect to its environment in the molecule. *That is*, one finds that the resonance

FIGURE 5
Mean Scores of University Registers Along Dimension 5,
Nonimpersonal Versus Impersonal Style



Note. $F = 117.6$; $df = 9, 453$; $r^2 = .70$; $p < .001$. The two figures in parentheses are mean scores and standard deviations, respectively.

frequency of an individual nucleus **is influenced** by the distribution of electrons in the chemical bonds of the molecule. (textbook: natural science [chemistry], graduate, tbchm3.gns)

The dense use of passives in textbooks serves as informational packaging. Noun phrases with the semantic role of agent or cause are less topically important than those with roles of patient or instrument; as a result, passive constructions are used to place the more important noun phrases in the grammatical subject position. The conjuncts explicitly mark the organization of the information and arguments.

Surprisingly, these constructions are also common in other campus

writing and in course management writing, which typically adopt an institutional rather than a personal voice. In these registers, references to students, the instructor, or the program administrator are often omitted, and the requirements, expectations, or other entities being influenced are fronted to the subject position, as illustrated by these excerpts from a department Web page:

the Master of Arts program in Anthropology is designed to . . .
when further information is required . . .
academic support is provided . . .
the groups are organized by the Center

and a course syllabus:

the order topics will be covered
students are expected to come to class prepared . . .
absences should be justified
your homework should be done with pride and submitted on time

Differences Across Disciplines and Levels

An analysis of dimension scores across disciplines and levels revealed some significant differences in textbooks but not in classroom teaching (see Appendixes A and B for descriptive statistics). ANOVAs (Table 5) showed significant differences for most dimensions among academic disciplines, within both classroom teaching and textbooks. However, these differences are generally not very strong, with r^2 values ranging from .06 to .36 (6–36%). Differences across levels are less marked, with all dimensions being nonsignificant except Dimension 5 for textbooks.

These findings, coupled with those described in the previous section, show considerable linguistic variation across university registers on the five dimensions and indicate that academic discipline and level are not associated with variation as much as register is. In fact, no significant variation was found among texts that differed in level, suggesting that students encounter generally the same structural linguistic features regardless of their level of study. As the preceding section has documented, texts in the various registers differ greatly in their linguistic features, but texts in the spoken and in the written modes show even greater differences. That is, regardless of specific purpose or subject matter, the physical mode of production seems to be by far the most important predictor of linguistic variation for university texts.

Obviously, the analysis reported here did not capture all linguistic

TABLE 5

Analysis of Variance for Classroom Teaching and Textbooks Across Disciplines and Levels

Dimension	<i>df</i>	<i>F</i>	<i>r</i> ²
Classroom teaching across disciplines (<i>n</i> = 176)			
1 ^a	5, 170	2.31	
2 ^b	5, 170	17.81*	.345
3 ^c	5, 170	13.17*	.279
4 ^d	5, 170	11.47*	.253
5 ^e	5, 170	1.53	
Textbooks across disciplines (<i>n</i> = 87)			
1 ^a	5, 81	2.51	
2 ^b	5, 81	9.09*	.362
3 ^c	5, 81	1.17	
4 ^d	5, 81	4.65*	.223
5 ^e	5, 81	6.77*	.295
Classroom teaching across levels (<i>n</i> = 176)			
1 ^a	2, 173	0.17	
2 ^b	2, 173	1.34	
3 ^c	2, 173	1.48	
4 ^d	2, 173	1.36	
5 ^e	2, 173	0.48	
Textbooks across levels (<i>n</i> = 87)			
1 ^a	2, 84	0.41	
2 ^b	2, 84	2.23	
3 ^c	2, 84	2.84	
4 ^d	2, 84	0.27	
5 ^e	2, 84	4.11	

Note. Probability was set at $p = .05$ and divided by 5 for each set of ANOVA to account for the use of multiple ANOVAs in each set of texts. The actual probability for determining significance was $p = .01$. No follow-up tests were conducted to test for differences among individual pairs because we were interested in the broad question of whether or not significant variation was identified.

^aInvolved versus informational production. ^bNarrative versus nonnarrative discourse. ^cSituation-dependent versus elaborated reference. ^dOvert expression of persuasion. ^eNonimpersonal versus impersonal style.

* $p < 0.01$.

differences across university texts. In particular, we expect more detailed investigations of vocabulary and the extent of assumed technical background knowledge to reveal important differences across disciplines and levels. These differences might be even sharper if considered across specific academic disciplines (e.g., biology, philosophy, sociology) rather than across macrodisciplines (e.g., humanities, natural sciences) as they are here. Despite these caveats, the MD analysis reported here shows a surprising leveling of linguistic form used in classroom teaching and textbooks, with few structural differences across disciplines and levels.

In a few minor exceptions to this generalization, however, disciplines or levels differ in their dimension scores. Because the differences are much smaller than those discussed in the preceding sections and because the specific disciplines/levels are less well represented than the more general registers are, we offer only tentative interpretations of those differences here. First, business classroom teaching is somewhat more interactive than the norm along Dimension 1, involved versus informational production; natural science classroom teaching is somewhat less interactive (see Appendix A). Education textbooks are also somewhat more involved than the norm for textbooks. These differences may reflect disciplines' preferred styles of instruction in teaching (i.e., class interaction vs. lecture style) and textbooks (i.e., a relatively interpersonal vs. a distanced relationship between the author and reader).

Along Dimension 2, narrative versus nonnarrative discourse, the humanities and education registers are somewhat more narrative than other disciplines are; this is true of both classroom teaching and textbooks. In the humanities, this pattern reflects the importance of historical recounts in subdisciplines like history, religious studies, and philosophy. Education seems to show a similar focus on narrative (either personal or historical).

The disciplinary differences along Dimension 3, situation-dependent versus elaborated reference, are more surprising, with classroom teaching in natural science and engineering (and, to a lesser extent, business) being considerably more situation dependent than the other disciplines. These patterns reflect the importance of physical demonstrations in the classroom teaching of those disciplines, with instructors repeatedly referring directly to displays or activities physically present in the classroom. Text Sample 11 illustrates teaching of this type in a computer science class in which the instructor refers to a computer display while discussing Visual Basic programming techniques. (Time and place adverbials and other adverbs are underscored.)

11. Instructor: OK what I wanted to do is another example, OK on the list box, let me try to get the right one here. OK and if you want to, so you don't have to keep up with your notes, you can make a copy of this, after class. OK.— So what I've got let me go ahead and run it, is a list of states, on the left side, and I want to display whichever one I've selected on the right side. So I take Colorado, and I push, this button, it takes Colorado off this side and adds it to this side. OK. If I select Colorado over here I can push that button and add it back. (classroom teaching: engineering [computer science], lower division, engcsleldln050)

In addition to the features of Dimension 3, personal pronouns (e.g., *I, you, we*) and demonstratives (e.g., *this, this side*) appear often in Text Sample 11. These features also exemplify frequent reference to the situation.

Surprisingly, engineering classroom teaching is also especially marked along Dimension 4, overt expression of persuasion, perhaps reflecting the same reliance on physical displays and demonstrations, and classes in which students are expected to consider alternative analyses and argue for a preferred solution. An example is the large number of conditional clauses and modal verbs used in Text Sample 12, an excerpt from the same classroom session as Text Sample 11. (Persuasive features are underscored.)

12. Instructor: Hey there's, actually while I'm thinking about it, there's also one other thing you might want to check here. What happens if I enter a current day that's less than the day I rented it? That's a bummer too right? OK. Think about it. In your head you need to think about all the possible mistakes that a user can make. OK one good way to do it is to get your kid or your next door neighbor to come over and try to break it . . . OK so you may also want to check if the sys- if, our day, OK. Now, when I have to do this text, then what do I need to do to this to make it usable? There's another function. OK it'll be (C.) day.—OK. Now I can take this, put it in there, OK, subtract what - this date? OK. If I rented, if this day is the same as the system day what's the answer? (classroom teaching: engineering [computer science], lower division, engcsleldln050)

Along Dimension 5, nonimpersonal versus impersonal style, engineering (and, to a lesser extent, natural science) is extremely marked for the dense use of passive constructions. This pattern fits the stereotypical characterization of technical and scientific prose. Interestingly, this difference exists only for textbooks; in contrast, we found no significant Dimension 5 differences among disciplines within lectures.

As we noted above, classroom teaching and textbooks almost never differed in dimension scores across levels (see Appendix B and Table 5). The sole exception to this generalization is the Dimension 5 differences for textbooks: Passive constructions are somewhat less common in lower-division than in upper-division and graduate-level texts. Thus, for the features studied here, the only concession in linguistic style made to entering undergraduates—in either classroom teaching or textbooks—is a less dense use of passive constructions in textbooks. Regardless of level, classroom teaching is relatively interactive and noninformational (Dimension 1), situated and not referentially elaborated (Dimension 3), and not passive (Dimension 5). In contrast, textbooks are consistently informational (Dimension 1) and referentially elaborated (Dimension 3), again regardless of level.

SUMMARY AND IMPLICATIONS

The findings of our multidimensional analysis of speaking and writing at the university have important implications for teaching and future research. Perhaps most important is the perspective gained on the range of language that students encounter at universities. On all dimensions, the university registers were found to cover a wide spectrum. On all dimensions except Dimension 2, narrative versus nonnarrative discourse, the corpus contained registers falling at both ends. Students must deal not only with informationally dense prose but also with interactive and involved spoken registers. They must handle texts with elaborated reference as well as those that rely on situated reference, and texts with features of overt persuasion as well as texts that lack those features. They must understand discourse that uses an impersonal style with many passives as well as discourse that tends to avoid passives. One of the noteworthy contributions of this study, therefore, is to begin to describe the linguistic challenge faced by students in U.S. universities. Teachers and researchers need to be aware that part of this challenge is students' need for facility in a tremendous range of registers.

The distribution of registers along Dimension 1, involved versus informational production, is particularly important. Academic registers are typically assumed to be extremely informational, but this study has shown that university students also encounter highly interactive, involved registers. Even registers with a strongly informational purpose, such as classroom teaching and study groups, are marked for the features of face-to-face interaction rather than the features of informational production. Previously, researchers and language teachers have paid little attention to the fact that students must rely on conversational language features to glean academic information from face-to-face interactions.

Another important finding of this study is that most dimensions show a strong polarization between spoken and written registers. The written registers—regardless of their specific purpose—are characterized by informationally dense prose, a very nonnarrative focus, elaborated reference, few features of overt persuasion, and an impersonal style. (The exception to this pattern is the course management register, which frequently shows features of overt argumentation.) In contrast, the spoken registers—again regardless of purpose—are characterized by features of involvement and interaction, situated reference, more overt persuasion, and fewer features of impersonal style. This finding contrasts with those of previous MD studies of English, which did not find spoken and written registers to be consistently polarized. For example, fiction writing is strikingly different from the written university registers considered here (see Biber, 1988, chapter 7). It falls near 0 on involved versus informational production (Dimension 1) and is marked strongly for the

use of narrative features (Dimension 2), situation-dependent reference (Dimension 3), and nonimpersonal style (Dimension 5). Students may well read fiction or other registers, such as newspapers, that have these relatively mixed profiles, but the oral and written university registers consistently differ in their features.

This division in the academic registers is especially surprising given the numerous purposes represented in the T2K-SWAL Corpus. The spoken registers, for example, range from interpersonal interactions with both social and informational purposes (e.g., service encounters and study groups), to monologic discourse with a primary informational focus (e.g., some types of classroom teaching). Students are regularly expected to integrate spoken and written material (Carson et al., 1992); the findings here suggest that this integration is likely to be challenging, given the polarization of linguistic characteristics across the modes.

Implications for Materials Development

This study has powerful implications for test development. The analysis describes the type of language that should inform such tests as the TOEFL if they are to accurately reflect the type of language used at universities. According to our results, students need the ability to handle not only academically dense prose but also interactive informational registers. In fact, Educational Testing Service is currently revising the TOEFL in part by using these data to check the consistency of test language with actual language use in university contexts (as represented in the T2K-SWAL corpus; see Educational Testing Service, 2001; Jamieson, Jones, Kirsch, Mosenthal, & Taylor, 2000).

Materials for teaching EAP also need to reflect knowledge about registers used at the university. Like the TOEFL, practice materials need to integrate patterns of language forms that are typically used for particular functions at the university. (For further discussion of this issue, see Byrd & Reid, 1997; Conrad, 2000.) Students need practice with the wide range of registers that they will encounter when they undertake university work. This study has shown that even registers meant to welcome and help students—such as other campus writing, which includes handbooks, catalogues, and informational Web pages—present information in dense, complicated syntactic structures. These kinds of texts can make useful practice materials, though they are rarely thought of as academic texts.

In addition to implications for testing and teaching, the results of this study also raise issues for university staff to consider. Most important is the finding that the register of other campus writing is extremely marked in its use of dense, informational prose. Most of the material in this

category is meant to help students navigate policies and procedures or attract students to programs. Program administrators and advisers obviously want students to understand this information, but such dense prose seems unlikely to facilitate students' understanding or attract them to programs. Less densely integrated prose would likely fit more closely the needs of the audience and the purpose of the texts.

Further Research

Although this study has revealed a great deal about the nature of language used at U.S. universities, more research is called for to expand the understanding of academic registers. For example, more detailed studies of specific disciplines might reveal similarities and differences across disciplines. Additional features—including rhetorical and lexical features—also deserve attention. In particular, vocabulary studies may uncover differences not identified in this MD analysis.

The way students respond to the diverse registers at the university also merits attention (cf. Carkin, 2001). For example, how do students deal with the contrast between the interactive discourse of the classroom and the informational prose of the textbooks and course packets? Similarly, studies of instructors' intentions would be valuable. Do instructors attempt to use interactional features of language to facilitate their instructional purpose in the classroom?

Although many questions about academic language remain, this study has made a substantial contribution to the description of academic discourse, providing a relatively comprehensive analysis of language use in the university. Our hope is that this analysis will be especially useful in increasing the TESOL field's understanding of the language tasks that students face when they enter a U.S. university.

ACKNOWLEDGMENTS

This project was supported by Educational Testing Service. We thank the numerous research assistants and student workers from California State University, Sacramento; Georgia State University; Iowa State University; and Northern Arizona University who helped collect, transcribe, scan, tag, tag-edit, and analyze the corpus.

THE AUTHORS

Douglas Biber is a Regents' Professor in the Applied Linguistics Program at Northern Arizona University. His research interests have focused on register variation, grammar and discourse, and corpus linguistics. He has published books on these topics with Cambridge University Press (1988, 1995, 1998), Oxford University Press (1994), and Longman/Pearson (1999, 2001).

Susan Conrad is an associate professor in the Department of Applied Linguistics at Portland State University. Her interests include English grammar, writing in the academic disciplines, and the application of corpus linguistic techniques within TESOL. Her most recent book (edited with Douglas Biber) is *Variation in English: Multi-Dimensional Studies* (Pearson Education).

Randi Reppen teaches in Northern Arizona University's MA-TESL and PhD in applied linguistics programs and directs the university's Program in Intensive English. Her interests include corpus linguistics and materials development. She is a coauthor of *Corpus Linguistics: Investigating Language Structure and Use* (Cambridge University Press).

Patricia Byrd is a professor in the Department of Applied Linguistics and ESL at Georgia State University, where she teaches graduate courses in English grammar and materials design. Her publications include scholarly work on English grammar and Web-based instruction along with ESL textbooks focused on grammar and academic writing.

Marie Helt is an assistant professor of applied linguistics at California State University, Sacramento, where she coordinates the graduate TESOL program. Her research interests are in corpus linguistics and applied sociolinguistics, and she is involved in the American Egyptian Master Teacher Exchange Program through the U.S. Agency for International Development.

REFERENCES

- 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.
- Atkinson, D. (1996). The philosophical transactions of the Royal Society of London, 1675–1975: A sociohistorical discourse analysis. *Language in Society*, 25, 333–371.
- Atkinson, D. (1999). *Scientific discourse in sociohistorical context: The philosophical transactions of the Royal Society of London, 1675–1975*. Mahwah, NJ: Erlbaum.
- Bazerman, C. (1988). *Shaping written knowledge: The genre and activity of the experimental article in science*. Madison: University of Wisconsin Press.
- Berkenkotter, C., & Huckin, T. (1995). *Genre knowledge in disciplinary communication*. Hillsdale, NJ: Erlbaum.
- Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Biber, D. (1994). An analytical framework for register studies. In D. Biber & E. Finegan (Eds.), *Sociolinguistic perspectives on register* (pp. 31–56). New York: Oxford University Press.
- Biber, D. (1995). *Dimensions of register variation: A cross-linguistic comparison*. Cambridge: Cambridge University Press.
- Biber, D., & Conrad, S. (2001). Quantitative corpus-based research: Much more than bean counting. *TESOL Quarterly*, 35, 331–336.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
- Biber, D., & Finegan, E. (1994). Intra-textual variation within medical research articles. In N. Oostdijk & P. de Haan (Eds.), *Corpus-based research into language* (pp. 201–222). Amsterdam: Rodopi. (Reprinted in *Variation in English: Multi-dimensional studies*, pp. 108–123, by S. Conrad & D. Biber, Eds., 2001, Harlow, England: Pearson Education)

- Biber, D., & Finegan, E. (1997). Diachronic relations among speech-based and written registers in English. In T. Nevalainen & L. Kahlas-Tarkka (Eds.), *To explain the present: Studies in the changing English language in honour of Matti Rissanen* (pp. 253–275). Helsinki, Finland: Société Néophilologique. (Reprinted in *Variation in English: Multi-dimensional studies*, pp. 66–83, by S. Conrad & D. Biber, Eds., 2001, Harlow, England: Pearson Education)
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *The Longman grammar of spoken and written English*. Harlow, England: Pearson Education.
- Biber, D., & Reppen, R. (in press). What does frequency have to do with grammar teaching? *Studies in Second Language Acquisition*.
- Byrd, P., & Reid, J. (1997). *Grammar in the composition classroom*. Boston: Heinle & Heinle.
- Carkin, S. (2001). *Pedagogic language in introductory classes: A multi-dimensional analysis of textbooks and lectures in biology and macroeconomics*. Unpublished doctoral dissertation, Northern Arizona University, Flagstaff.
- Carson, J., Chase, N., Gibson, S., & Hargrove, M. (1992). Literacy demands of the undergraduate curriculum. *Reading Research and Instruction*, 31, 25–50.
- Cazden, C. (1988). *Classroom discourse: The language of teaching and learning*. Portsmouth, NH: Heinemann.
- Chaudron, C., & Richards, J. (1986). The effect of discourse markers on the comprehension of lectures. *Applied Linguistics*, 7, 113–127.
- Conrad, S. (1996). Investigating academic texts with corpus-based techniques: An example from biology. *Linguistics and Education*, 8, 299–326.
- Conrad, S. (1999). The importance of corpus-based research for language teachers. *System*, 27, 1–18.
- Conrad, S. (2000). Will corpus linguistics revolutionize grammar teaching in the 21st century? *TESOL Quarterly*, 34, 548–560.
- Conrad, S. (2001). Variation among disciplinary texts: A comparison of textbooks and journal articles in biology and history. In S. Conrad & D. Biber (Eds.), *Variation in English: Multi-dimensional studies* (pp. 94–107). Harlow, England: Pearson Education.
- Conrad, S., & Biber, D. (Eds.). (2001). *Variation in English: Multi-dimensional studies*. Harlow, England: Pearson Education.
- Crompton, P. (1997). Hedging in academic writing: Some theoretical problems. *English for Specific Purposes*, 16, 271–287.
- Csomas, E. (2000). Academic lectures: An interface of an oral and literate continuum. *novELT*, 7(3), 30–46.
- Cutting, J. (1999). The grammar of the in-group code. *Applied Linguistics*, 20, 179–202.
- Educational Testing Service. (2001). *Enhancing the standard with a new TOEFL test*. Princeton, NJ: Educational Testing Service.
- Edwards, J. A., & Lampert, M. D. (Eds.). (1993). *Talking data: Transcription and coding in discourse research*. Hillsdale, NJ: Erlbaum.
- Ervin-Tripp, S. (1972). On sociolinguistic rules: Alternation and co-occurrence. In J. J. Gumperz & D. Hymes (Eds.), *Directions in sociolinguistics* (pp. 213–250). New York: Holt, Rinehart & Winston.
- Ferguson, G. (2000). If you pop over there: A corpus-based study of conditionals in medical discourse. *English for Specific Purposes*, 20, 61–82.
- Flowerdew, J. (Ed.). (1994). *Academic listening: Research perspectives*. New York: Cambridge University Press.
- Flowerdew, J., & Tauroza, S. (1995). The effect of discourse markers of second language lecture comprehension. *Studies in Second Language Acquisition*, 17, 435–458.
- Gilbert, G., & Mulkay, M. (1984). *Opening Pandora's box: A sociological analysis of scientific discourse*. Cambridge: Cambridge University Press.

- Grabe, W., & Kaplan, R. B. (1996). *Theory and practice of writing*. London: Longman.
- Grabe, W., & Kaplan, R. B. (1997). On the writing of science and the science of writing: Hedging in scientific text and elsewhere. In R. Markkanen & H. Schroeder (Eds.), *Hedging in discourse* (pp. 151–167). Berlin: de Gruyter.
- Halliday, M. A. K. (1988). On the language of physical science. In M. Ghadessy (Ed.), *Registers of written English* (pp. 162–178). London: Pinter.
- Halliday, M. A. K., & Martin, J. R. (1993). *Writing science: Literacy and discursive power*. Pittsburgh, PA: University of Pittsburgh Press.
- Helt, M. (2001). A multi-dimensional comparison of British and American spoken English. In S. Conrad & D. Biber (Eds.), *Variation in English: Multi-dimensional studies* (pp. 171–184). Harlow, England: Pearson.
- Holmes, J. (1988). Doubt and certainty in ESL textbooks. *Applied Linguistics*, 9, 21–43.
- Huckin, T., & Pesante, L. H. (1988). Existential *there*. *Written Communication*, 5, 368–391.
- Hunston, S. (1995). A corpus study of some English verbs of attribution. *Functions of Language*, 2, 133–158.
- Hyland, K. (1994). Hedging in academic writing and EAP textbooks. *English for Specific Purposes*, 13, 239–256.
- Hyland, K. (1996a). Talking to the academy: Forms of hedging in science research articles. *Written Communication*, 13, 251–281.
- Hyland, K. (1996b). Writing without conviction? Hedging in science research articles. *Applied Linguistics*, 17, 433–454.
- Jamieson, J., Jones, S. Kirsch, I., Mosenthal, P., & Taylor, C. (2000). *TOEFL 2000 framework: A working paper* (TOEFL Monograph Series MS-16). Princeton, NJ: Educational Testing Service.
- Khuwaileh, A. A. (1999). The role of chunks, phrases, and body language in understanding coordinated academic lectures. *System*, 27, 249–260.
- Kuo, C.-H. (1999). The use of personal pronouns: Role relationships in scientific journal articles. *English for Specific Purposes*, 18, 121–138.
- Love, A. (1993). Lexico-grammatical features of geology textbooks: Process and product revisited. *English for Specific Purposes*, 12, 197–218.
- Marco, M. J. L. (1999). Procedural vocabulary: Lexical signaling of conceptual relations in discourse. *Applied Linguistics*, 20, 1–21.
- Marco, M. J. L. (2000). Collocational frameworks in medical research papers: A genre-based study. *English for Specific Purposes*, 19, 63–86.
- Myers, G. (1989). The pragmatics of politeness in scientific articles. *Applied Linguistics*, 10, 1–35.
- Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. Oxford: Oxford University Press.
- Reppen, R. (2001). Register variation in student and adult speech and writing. In S. Conrad & D. Biber (Eds.), *Variation in English: Multi-dimensional studies* (pp. 187–199). Harlow, England: Pearson.
- Salager-Meyer, F. (1999). Referential behavior in scientific writing: A diachronic study (1810–1995). *English for Specific Purposes*, 18, 279–305.
- Strodt-Lopez, B. (1991). Tying it all in: Asides in university lectures. *Applied Linguistics*, 12, 117–140.
- Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Swales, J. M., Ahmad, U. K., Chang, Y. Y., Chavez, D., Dressen, D. F., & Seymour, R. (1998). Consider this: The role of imperatives in scholarly writing. *Applied Linguistics*, 19, 97–121.
- Thompson, G., & Ye, Y. (1991). Evaluation in the reporting verbs used in academic papers. *Applied Linguistics*, 12, 365–382.

- Valle, E. (1999). *A collective intelligence: The life sciences in the Royal Society as a scientific discourse community, 1665–1965*. Unpublished doctoral dissertation, University of Turku, Finland.
- Varantola, K. (1984). *On noun phrase structures in engineering English*. Unpublished doctoral dissertation, University of Turku, Finland.
- Williams, I. (1996). A contextual study of lexical verbs in two types of medical research report: Clinical and Experimental. *English for Specific Purposes*, 15, 175–197.

APPENDIX A

Descriptive Statistics for Classroom Teaching and Textbooks by Discipline

Dimension	Discipline	Classroom teaching			Textbooks		
		Texts	<i>M</i>	<i>SD</i>	Texts	<i>M</i>	<i>SD</i>
1 ^a	Business	36	33.3958	11.4157	15	-15.3192	6.5273
	Education	16	29.5925	8.9060	6	-10.5583	7.2675
	Engineering	30	29.5393	10.3706	9	-16.9389	2.5622
	Humanities	31	27.5781	10.2165	18	-16.2339	8.2272
	Natural sciences	25	24.9760	10.0419	18	-18.8972	3.8659
	Social sciences	38	27.3711	10.8059	21	-18.7825	5.1699
2 ^b	Business	36	-1.4958	1.3813	15	-3.1800	1.0388
	Education	16	-0.7369	0.8385	6	-1.1250	3.0422
	Engineering	30	-2.3480	1.0175	9	-4.4322	0.6704
	Humanities	31	-0.6587	1.0128	18	-1.6700	1.7250
	Natural sciences	25	-2.4680	0.5660	18	-3.8228	1.0047
	Social sciences	38	-0.8908	0.8244	21	-2.9217	0.9164
3 ^c	Business	36	3.6425	1.8751	15	-6.5700	2.2421
	Education	16	1.6444	2.0084	6	-4.8983	4.1576
	Engineering	30	4.5703	2.5611	9	-5.9456	1.8311
	Humanities	31	1.4942	2.2867	18	-6.3033	3.4560
	Natural sciences	25	4.6684	1.3425	18	-4.8089	2.6598
	Social sciences	38	1.6487	2.7049	21	-6.5446	2.4460
4 ^d	Business	36	2.7911	1.7362	15	0.1408	2.1345
	Education	16	1.2381	1.7950	6	-0.4833	2.0079
	Engineering	30	4.2827	2.3973	9	-1.3589	1.2365
	Humanities	31	1.2784	2.5883	18	-1.8256	2.7293
	Natural sciences	25	1.8976	2.1090	18	-2.7294	1.4128
	Social sciences	38	0.7545	2.0127	21	-2.5704	1.6233
5 ^e	Business	36	1.3447	0.8493	15	-3.8733	2.3406
	Education	16	1.3919	0.5031	6	-2.0833	2.1484
	Engineering	30	1.2783	1.0541	9	-6.3233	1.5396
	Humanities	31	1.0623	0.7550	18	-2.3533	1.3637
	Natural sciences	25	1.2172	0.8341	18	-4.7789	2.2460
	Social sciences	38	0.8526	1.1470	21	-3.8033	1.9856

^aInvolved versus informational production. ^bNarrative versus nonnarrative discourse. ^cSituation-dependent versus elaborated reference. ^dOvert expression of persuasion. ^eNonimpersonal versus impersonal style.

APPENDIX B

Descriptive Statistics for Classroom Teaching and Textbooks by Level

Dimension	Level	Classroom teaching			Textbooks		
		Texts	<i>M</i>	<i>SD</i>	Texts	<i>M</i>	<i>SD</i>
1 ^a	Lower division	54	28.8441	11.5254	28	-16.1932	6.6396
	Upper division	72	28.4128	11.0915	28	-17.2846	5.8793
	Graduate	50	29.5606	9.2620	31	-17.5932	6.0119
2 ^b	Lower division	54	-1.2085	1.2559	28	-2.4504	1.8685
	Upper division	72	-1.5528	1.1721	28	-2.9136	1.7079
	Graduate	50	-1.4998	1.2403	31	-3.3419	1.2597
3 ^c	Lower division	54	3.4496	2.7835	28	-5.3239	2.7028
	Upper division	72	2.8024	2.6605	28	-5.5700	2.7894
	Graduate	50	2.6436	2.2079	31	-6.8968	2.7371
4 ^d	Lower division	54	2.3013	2.5059	28	-1.9543	2.0862
	Upper division	72	1.7079	2.2187	28	-1.5614	2.3516
	Graduate	50	2.3454	2.6430	31	-1.8929	2.0124
5 ^e	Lower division	54	1.1372	0.9298	28	-2.9179	2.3578
	Upper division	72	1.1092	0.9619	28	-4.0746	1.8724
	Graduate	50	1.2704	0.8792	31	-4.5084	2.2811

^aInvolvement versus informational production. ^bNarrative versus nonnarrative discourse. ^cSituation-dependent versus elaborated reference. ^dOvert expression of persuasion. ^eNonimpersonal versus impersonal style.