11. Linguistic Variation in the Lexical Episodes of University Classroom Talk

11

Linguistic Variation in the Lexical Episodes of University Classroom Talk

ENIKO CSOMAY
San Diego State University

Background

The linguistic characteristics of texts have been researched from two major perspectives over the past three decades: one describing the internal discourse organization of texts and the other focusing on the typical linguistic characteristics of texts and text types. Studies of the first type usually have been qualitative, providing detailed analyses of the discourse patterns in individual texts (e.g., Fox 1987; Mann, Matthieson, and Thompson 1992). In contrast, studies of the second type provide quantitative results, using lexical and grammatical features for their analyses while generally ignoring higher-level discourse structures (e.g., Biber 1988, 1995; Reppen, Fitzmaurice, and Biber 2002).

The linguistic studies analyzing university classroom discourse reflect these two major research traditions. On one hand, researchers investigate the discourse patterns of academic lectures by focusing on lexical patterns and their relationship either to coherence (Tyler 1995) or to the micro and macro structures of class sessions (DeCarrico and Nattinger 1988). Other studies concentrate on the subunits in lectures and identify the linguistic markers of topic shifts (e.g., Hansen 1994) or relate the subunits to various communicative purposes (Young 1994) in the classroom. Although these studies are invaluable in our understanding of the discourse patterns in lectures, most of the analyses are carried out on a few selected texts; hence, the findings are not generalizable.

On the other hand, most recent studies have provided us with comprehensive linguistic characterizations of university classroom discourse, relying on quantitative analysis of a wide range of co-occurring linguistic features. Using Biber's (1988) dimensions of textual variation across registers, Csomay (2000) concludes that university classrooms exhibit linguistic features of both academic prose and face-to-face conversation. In describing the dimensions of textual variation within the academic context, Biber (2003) compares the linguistic characteristics of classrooms to language used in other academic registers (e.g., textbooks). Finally, Csomay (2002a) investigates linguistic variation across class sessions and describes variation in language use related to the level of instruction and the degree of interactivity (i.e., monologic versus interactive classes). Although these studies are able to draw

generalizations concerning patterns of linguistic variation across texts, they fail to show variation within texts.

My goal in this chapter is to provide a linguistic characterization of lexically coherent discourse units found in university classroom texts, using quantitative measures for the analysis. The discourse units characterized this way correspond to varying communicative purposes and provide the foundation to further describe text-internal linguistic variation.

More specifically, I first apply Youmans's (1991) Vocabulary Management Profile (VMP) to trace the overall lexical patterns in discourse. On the basis of this lexical profile and a quantitative segmentation technique developed for the study, I identify two major discourse units: lexical episodes and transitional units. Corresponding to lexical episodes is the introduction of new lexis in a stretch of discourse; primary reliance on repeated vocabulary marks transitional talk between the lexical episodes. Second, I classify lexical episodes into types on the basis of their shared linguistic characteristics, using Biber's (2003) dimensions of academic discourse. Finally, I associate these lexical episode types (involved narrative, procedural, content-oriented) with varying communicative purposes and present them in a partial taxonomy.

Analytical Procedures

I took three major analytical steps to achieve the main goals of the study: segment class sessions into discourse units; identify multidimensional characteristics of lexical episodes; and identify lexical episode types on the basis of their multidimensional characteristics of academic discourse.

Segment Class Sessions into Discourse Units—The VMP

The VMP tracks the introduction of newly occurring vocabulary into the discourse. To identify newly occurring vocabulary items, the text is processed via a "sliding window" of 100 words, and new vocabulary items are counted in that window at each word. That is, at the start the window is positioned at the beginning of the text and contains words 1–100. Then the window "slides" one position and contains words 2–101. The window continues to slide one position at a time, allowing analysis of overlapping 100-word chunks of the target text, until the end of the text is reached.

The VMP keeps track of every word introduced in the text prior to the point of analysis, while keeping track of the "new" words entering the window. A word counts as new if it had not occurred in the text before. Thus, a word could occur only once in the window but *not* be counted as new because it had been used earlier in the text. The value of the VMP can change plus or minus 1 every time the window slides one position. Because we use a 100-word window for the analysis, the VMP has a potential range of 0–100. That is, at each position of the window there can be anywhere from 0 to 100 new words occurring in that 100-word chunk of text. In practice, however, the VMP ranges from 5 to 40.

Using the method described above, we gain a VMP value for each word sliding through the window. If we plot the VMP values for each word, we obtain the visual representation of the dynamic change in alternating repeated and newly occurring vocabulary items in a discourse. Figure 11.1 shows the VMP for a 1,500-word

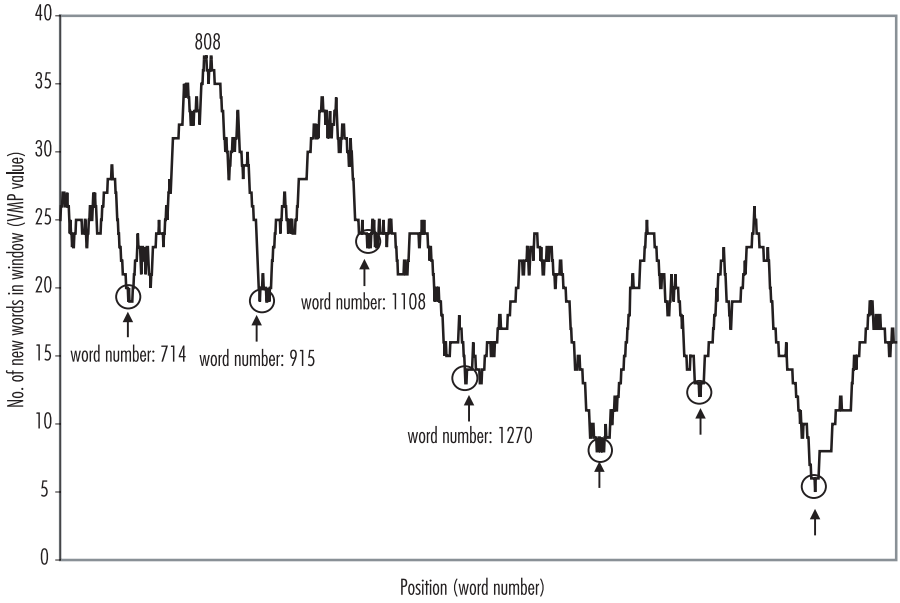

Figure 11.1. VMP patterns for a 1,500-word text segment in university classroom discourse

segment from a university classroom text (from word position 500 to word position 2,000), where each data point represents the number of new lexical items in the 100-word window at that point in the text.

Showing the patterns of "old" and "new" vocabulary in the discourse this way, we can identify the vocabulary patterns in a text. A steady introduction of new vocabulary items is associated with the beginning of a lexical episode (e.g., 714–808 in figure 11.1). A steady decline in the VMP (as in positions 1,108–1,270 in figure 11.1) represents recycling of old vocabulary and is interpreted as the coda of a lexical episode. The former is called a "lexical episode"; the latter is called a "transitional unit." Extract 1 illustrates the onset of a lexical episode (positions 714–808 in figure 11.1); the new words entering the discourse are capitalized.

Extract 1: Philosophy

Teacher: . . . What do they say that is relevant to the question of how we (714) ought to behave? We can look for the ANSWERS to THOSE questions, we will be TRYING to DEVELOP an EXISTENTIALIST ethic. It's PRETTY CLEAR that Sartre . . . and I'll be TALKING PRIMARILY about Sartre, REJECTS most of the STANDARD FORMS of ethics that you encounter in the HISTORY of WESTERN PHILOSOPHY. Not all of these ethical THEORIES FOCUS on RULES and PRINCIPLES, but most of them do. Most ethical theories TRY to ASCERTAIN the FUNDAMENTAL principles of MORALITY. JOHN STEWART LILL in the NINETEENTH CENTURY DID EXACTLY that, he said what is the (808) most basic and fundamental principle of right and wrong.

After these textual measures were developed, a two-step computational tool was developed to automatically recognize lexical episodes on the basis of the numeric VMP values computed in the previous step. In this study, lexical episodes are identified through three points: the peak and the two surrounding valley points in the VMP value. First the important VMP peaks were identified. "Important" peaks represent a high point in the VMP following a sustained increase from the preceding valley—at least a ten-point difference in the VMP between a peak and a valley point (i.e., ten new words used in the overall discourse).

Second, the boundaries surrounding the important peaks were identified through the sliding slope measures of a best-fitting regression line. The slope of the regression line could be closest to 0.0 only at the deepest valley points on the two sides of the peaks (for a more detailed description, see Csomay 2002b).

The final step in preparing for the linguistic analysis was to build a corpus of lexical episodes. The classroom texts were taken from the Test of English as a Foreign Language (TOEFL) 2000 Spoken and Written Academic Language (T2K-SWAL) Corpus. This corpus (about 2.7 million words) was designed to represent the language used in the different registers associated with the academic setting in universities of the United States (Biber et al. 2002). Using this corpus and applying the methodology described above, a total of 2,200 lexical episodes were extracted from 176 (about 1.2 million words) university classroom teaching texts.

Identify Multi-Dimensional Characteristics of Episodes

The second step in the analytical procedures was to document the linguistic characteristics of lexical episodes. The multidimensional analysis developed by Biber (1988) shows parameters of linguistic characteristics that work together. In his latest study, Biber (2003) counted more than ninety linguistic features (cf. Biber et al. 1999) to identify four underlying dimensions characteristic to academic discourse: oral versus literate discourse, procedural versus content-focused discourse, narrative versus nonnarrative orientation, and academic stance. The linguistic features identified on these four dimensions (table 11.1) were used to characterize lexical episodes in this study.

Identify Lexical Episode Types on the Basis of Their Linguistic Characteristics

Lexical episode types are identified on the basis of their linguistic characteristics on the four dimensions. To identify groupings of episode types, statistical measures—cluster analytical methods—are used. In applying cluster analysis, the ultimate goal was to find linguistic similarities in lexical episodes rather than to focus on the linguistic differences they may exhibit. That is, through statistical methods, lexical episodes with similar linguistic characteristics group together into a cluster reflecting shared communicative purposes. At the same time, the different clusters interpreted as lexical episode types distinctly "differ from one another in that they have different linguistic characterizations and correspondingly different functional emphases" (Biber 1995, 321).

Table 11.1.
Selected linguistic features on the four dimensions of academic language

Dimensions	Negative features	Positive features
1	Literate discourse	Oral discourse
	Prepositions	THAT omission
	Attributive adjectives	Demonstrative pronoun
	Passives	Present tense
	Nouns (group, human, mental)	Adverbials (time, place, hedges)
	Causative verbs	1st, 2nd, 3rd person pronoun
		Discourse particles
2	Content-focused discourse	Procedural discourse
	Content words in one-text-only	2nd person pronoun
	Size attrib. adj.	Modals (predictive and necessity)
	Past tense	Nouns (group, moderately common)
	Moderately common adverbs	Verbs (moderately common, activity, causative)
	Occurrence verbs	TO-clause controlled by Desire verbs
	BY passives	
3	Non-narrative orientation	Narrative orientation
	Noun (technical concrete, concrete, quantity)	Past tense
		3rd person pronoun
		Nouns (human, mental)
		Verbs (mental, common, communication)
		THAT omission
4	Wh- Questions	Academic stance
	WH- questions	Adverbials (factual)
	Stranded prepositions	THAT clause with Noun
		Adverb (likelihood, attitudinal)

Source: Biber (2003).

Findings

Based on the four dimensions of academic discourse (see foregoing discussion), three clusters were identified. The three clusters were interpreted as three lexical episode types, which were then examined for their communicative purposes.

Lexical Episode Types

Three major lexical episode types were identified: involved narrative, procedural, and content-oriented. Figure 11.2 summarizes the characteristics of each episode type with respect to the mean dimension scores on three of the four dimensions of academic language.

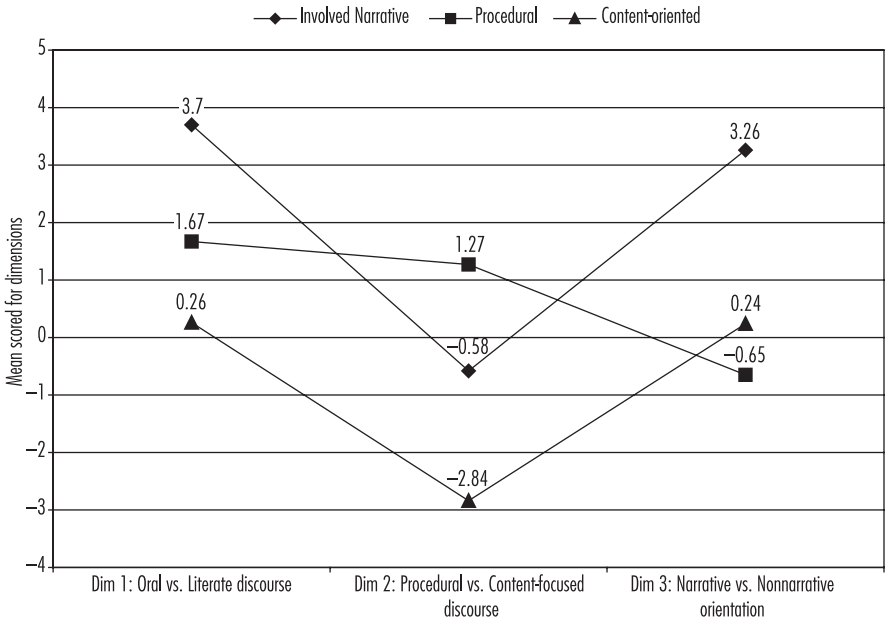

Figure 11.2. Three lexical episode types on three dimensions of academic language

As figure 11.2 shows, involved narrative episode types show the highest positive scores on two dimensions: oral discourse and narrative orientation. Procedural episodes show the highest scores on procedural discourse, and content-oriented episodes exhibit linguistic features associated with literate and content-focused discourse.

Involved Narrative Lexical Episodes The first lexical episode type, involved narrative, is distinguished by the high frequency counts of linguistic features on two dimensions: the positive side of dimension 1 and dimension 3. Positive dimension 1 features include, for example, first-, second-, and third-person pronouns, contractions, non–past tense, demonstrative pronouns, and so on—reflecting an involved, interactive style and showing personal stance. Positive dimension 3 features include past tense, animate and cognitive nouns (e.g., knowledge, fact, understanding), mental verbs (e.g., know, think, believe) (see Biber et al. 1999), and verbs commonly used in the T2KSWAL corpus as a whole (e.g., be, think, go). Extract 2 shows part of a text segment exhibiting these features.

Extract 2: Humanities

Teacher: yeah. so, I I I would agree that that B. doesn't have a terribly effective presentation style.

Student: especially when he's sick

Teacher: no presence

Student: no, he does not.

Student: well and it was interesting to me that this sort of big deal public lecture thing seemed to me to be the least well thought out and sort of coherent, I mean the, the forty five minute, or ya know forty minute whatever he did in here was absolutely stunning. ya know, ya know and I

Teacher: I have marvelously organized notes based on what he was saying

Student: yeah, and I felt like he, um had thought about what he was gonna say

Student: but don't you think this is his round,

Student: could be, could be

Student: the classroom is his world. I mean standing up in front in the pioneer room

Student: yeah, that's true, after meeting him it is less it it it did seem less of his, uh domain.

Student: yeah

Student: because he is a very good speaker just sitting with him at lunch and asking him questions and

Student: right

Student: ya, know

Student: plus he's sick to death of medicine, I mean he's not interested in talking about medicine.

Student: every form of writing, every [unclear statement]

Student: well I have [unclear statement] being in the mood for medicine. I mean he said this was eighteen months ago

Student: every time he talked about it, he made that [unclear]

Student: he said that it might as well been a lifetime ago, he said, he said, I'm more interested in talking about where I am now and these researches that I'm doing. I mean he's been away from medicine for quite awhile and now he went back . . .

The foregoing extract shows a lexical episode in which the grammatical features on the positive side of dimension 3 and positive dimension 1 are present.

With regard to the functional interpretation of this lexical episode type, involved narrative episodes were associated with situations in which information related to personal experiences was shared or concepts were explained and interpreted through relating text or visual input to personal experiences, feelings, and beliefs. Alternatively, this lexical episode type occurred in phases in which the teacher elaborated on or summarized known information.

Procedural Lexical Episodes The next episode type, procedural episodes, exhibited linguistic features on the positive side of dimension 2—for example, predictive modals (e.g., will), desire verbs with *to* clauses (e.g., want), second person pronouns (you, your), or nouns occurring moderately frequently in the entire corpus (e.g., chapter, exam, week, note, word, fact). These features are displayed in a high number in

extract 3. Linguistic features on the negative side of dimension 3, such as (technical) concrete nouns (e.g., book, case, formula, exam) or quantity nouns (e.g., week, today) also appear frequently in this segment.

Extract 3: Engineering

Teacher: the medium high is is is about (eight at eighteen) hundred and then the other is twenty four hundred ok that. We'll try it again if you still have problems talk to me after class anybody else have comments or questions? Ok . . . um . . . a couple of things just remind you in *case* you are not . . . in *case* you haven't just kept a quick eye on the *syllabus* that we have the exam scheduled a *week* from *today* . . . *uh exam* one covers through the first two *chapters* so I wanted to bring that to your attention um just in your preparation between now and then uh if you look at the problems or going through your *book* or whatever um the *exam* will be closed *book* but you can bring in a *sheet* of *notes* in other words you can make a legal cheat *sheet*. . . uh and you and the thought process there is I don't want you to spend all *exam* time thumbing through your *book* looking for a formula to match up or an example problem to match up but I don't also don't want you to spend your time memorizing the *formula* and worrying whether you properly memorized a *formula* when in fact all I want you [unclear words] be able to use it so.. in your *notes* . . .</EXT>

Extract 3 shows a lexical episode in which the grammatical features on the positive side of dimension 2 and negative dimension 3 are exhibited.

Procedural episodes reflect instructional and/or communicative purposes such as sharing information related to either the classroom context (e.g., how to go about the midterm test) or professional experiences (e.g., what kinds of activities students will be engaged in once they leave the program), presenting the methods of a classroom project, or explaining concepts step by step while relating visual input to prior knowledge (e.g., interpreting a graph).

Content-Oriented Lexical Episodes The third lexical episode type, content-oriented episodes, relied on linguistic features such as prepositional phrases, relative clause constructions (reflecting elaboration), past tense, or content words that occurred in one class session only. Extract 4 is an example of a content-oriented episode.

Extract 4: Humanities

Student: . . . By eighteen twenty, Americans had **moved** INTO another century, not only IN time but IN thought. IN the way they perceive themselves IN the world. They had **experienced** the social and cultural transformation as great as any IN American history. The transformation MARKED **BY** the search FOR an American identity and **BY** the climax and fall of the enlightenment IN America. The American Revolution **seems to** present Americans WITH an opportunity to realize an ideal world. To put the enlightenment INTO practice, to create the **kind** OF ordered society and illustrious CULTURE

> THAT man since the Greeks have **yearned** FOR. WITH the revolution and the IDEAS IN enlightenment that accompany to contain WITHIN themselves the SOURCES OF their own disillusionment and destruction. BY eighteen twenty the enlightenment OF America **was** over. The ideals OF the revolution **changed** and **perverted**, yet the transformation **was** so complicated, so **undeliberate**, so [unclear], so much a **medley**, OF responses [unclear] events that Americans scarcely **knew** how they **got** FROM one point TO the other. They **began** the revolution BY then IN a world they **felt** very much a part OF. They **ended** by perceiving their destiny IN America itself, BY becoming a peculiar and unprecedented **kind** OF republic. It was so - it was an **unattended** revolution FOR the character they **saw** revealed IN Andrew Jackson and the hundreds OF Kentucky romantic, undisciplined, and untutored heroes OF the battle OF New Orleans **was** scarcely the character they **saw** IN seventeen seventy six. The [unclear] nationalism [unclear] AT the end OF the war OF eighteen twelve. It **represented** both a repudiation OF the classic ideals OF the REVOLUTION and an attempt to come to terms WITH largely unanticipated society that **emerged** FROM the revolution. A new CULTURE that had been **created** both because and IN spite of the REVOLUTION. WITH the peace OF Ghent and the end OF the Napoleonic WARS the new American republic seemed at last secure and ready to comprehend itself.

The relatively few linguistic features from positive dimension 1 are shown through the relatively few cases of underlined words (non–past tense; first-, second-, third-person pronouns; etc.). The linguistic features from negative dimension 2 in this segment are in ***bold italics*** (past tense, *by* passives, content words occurring in one text only, etc.). In addition, some features (e.g., prepositions, relative clauses, and some nouns) on the negative pole of dimension 1 are highlighted by capital letters.

This episode type was associated with purposes that reflected a strong informational focus. They occurred when a monologue-like teaching style was adopted or texts were read out loud (as in extract 4). Content-oriented episodes also reflected situations in which the explanation of a particular concept was supported by visual aids (e.g., through writing on the board).

Partial Taxonomy of Lexical Episode Types and Their Communicative Purposes

As the foregoing examples show, lexical episode types sharing similar linguistic characteristics have similar communicative purposes. The communicative tasks identified in the lexical episodes as found in university classroom discourse are shown in table 11.2.

The communicative purposes listed in this partial taxonomy all relate to information. This finding may not be surprising. The present research set out to analyze units in university classroom discourse that were marked for vocabulary items newly entering the discourse and were interpreted to largely correspond to new topics and new information occurring in the stretch of discourse. In most cases, each lexical

Table 11.2.
Partial taxonomy of communicative tasks as found in lexical episodes of university classroom talk

Communicative Purpose	Episode Type		
1. Share information on			
a. classroom procedures	P		IN
b. personal experiences			IN
c. professional experiences	P		
d. study skills	P		
2. Transmit information through			
a. lecturing		C	
b. reading a text aloud		C	
3. Interpret/explain information through relating			
a. text or visual input to			
• prior knowledge	P		
• personal experiences, feelings, beliefs			IN
• new information in the class session		C	
b. conceptual information to			
• visual illustration in context (e.g., board in the classroom, diagram in the book)		C	
4. Summarize/elaborate on information expected as known			IN
5. Report on project(s) by			
a. student	P		
b. teacher			IN
6. Demonstrate how something works (e.g., a computer program)	P		

Note: IN = involved narrative; C = content-oriented; P = procedural.

episode has a clear topical focus and a single communicative purpose. Although each episode is lexically coherent, there is variation in the way information is conveyed in these episodes.

Based on the partial taxonomy presented here, we can see what kinds of communicative tasks the different lexical episode types are most likely to exhibit. For example, there is a clear distinction between content-oriented episodes (information transmitted via lecturing or via reading a text out aloud) and procedural or involved narrative episodes (information shared through personal anecdotes or through talking about professional experiences). A most interesting point about the variation in the way information is conveyed comes out when information is related to textual or visual input (point 3 in table 11.2). Although content-oriented lexical episodes use visuals (either textual or figurative) to support new information (e.g., writing on the board), both procedural and involved narrative episodes use visuals (textual as well as figurative) as primary sources for presenting new information, based on which

prior knowledge or personal aspects of the theme are connected. This distinction suggests a fundamentally different approach to conveying information in the classroom.

Summary and Future Directions

The primary goal of the present study was to provide a linguistic characterization of lexically coherent discourse units found in university classroom texts, using quantitative measures for the analysis. Based on the linguistic patterns in the discourse units, three major lexical episode types were identified (involved narrative, procedural, and content-oriented), each exhibiting varying communicative purposes in the thread of discourse.

In its approach, the present study complements earlier studies in three areas: linguistic analyses of discourse patterns in university classroom talk; corpus-based analyses of discourse units and textual variation; and identification of discourse units in texts. First, the research described here contributes to the analysis of classroom discourse patterns in which discourse units are identified and assessed through vocabulary novelty that participants experience in the discourse event of a class session. This approach offers a new perspective to analyzing classroom discourse patterns not only in the university setting; it also could be extended to other instructional settings (e.g., English as a Second Language/English as a Foreign Language classrooms).

Second, the lexically coherent discourse units were analyzed by using corpus-based analytical techniques, providing a comprehensive linguistic analysis of those units. By segmenting discourse using reliable measures and by developing a taxonomy of lexical episode types on the basis of their lexicogrammatical characteristics, we can establish the foundations for describing patterns of text-internal variation.

Third, empirical methods were developed to identify discourse units in (classroom) discourse, providing replicable findings of the present research. Although the discourse units have been predefined on the basis of their perceived communicative and instructional purposes in earlier classroom-based studies, the present research recognized those purposes after the units had been identified and characterized on the basis of reliable measures.

Although the study provides new perspectives for analyzing classroom discourse and points to future directions in this area, it also has limitations. First, the methodology could be improved for a more precise identification of topical units in discourse. Although the VMP tracks the introduction of new vocabulary items into a discourse, which often marks new topics in the unfolding discourse, vocabulary novelty is identified relative to the discourse stretch occurring prior to the point of analysis. Hence, this methodology fails to identify a lexical segment as "important"—that is, denoting a new topic—if recycled vocabulary is used, leaving out the possibility of new topics worded with recycled vocabulary in the same stretch of discourse. This pattern may be an important characteristic of classroom discourse, however, because class participants may approach different topics while verbalizing those topics with words they already have used once during that session. Applying a modified version of Hearst's (1997) TextTiling, an alternative methodology to identify topical units with greater precision is suggested by Biber et al. (2004). Second, more discourse units could be

analyzed. In this analysis, only lexical episodes were included in the linguistic analysis. The discourse units that use recycled vocabulary and serve as links between lexical episodes were called "transitional units." These units were not included in the present linguistic analysis; they are equally important, however, in overall discourse patterns. Compared to lexical episodes, they may be very similar or, indeed, very different in their linguistic characteristics and their communicative functions.

Finally, another limitation is related to psycholinguistic and pedagogical issues. This research is unable to show the cognitive difficulty students may encounter while they are exposed to or involved in any of the lexical episode types (although the study did not set out to investigate this area). Moreover, because of the lack of data measuring student performance, no conclusions can be drawn about whether classes exhibiting episodic patterns of particular lexical episode types are instructionally more effective than classes exhibiting another pattern. However, the three lexical episode types displayed differences in the way new information was linked to the actual textual input or to the visuals in class. This finding provides complementary linguistic evidence on classroom literacy events (Poole 2003), and from a teacher educational perspective it is a particularly important area of research. The results could provide teachers with evidence of how language is used in the different communicative tasks performed in the classroom.

In conclusion, descriptive studies of this kind are useful and could complement other classroom-based research because they offer replicable findings, rely on data collected from naturalistic settings, and provide data for the linguistic characterizations of what actually happens in a large number of classrooms.

REFERENCES

Biber, Douglas. 1988. *Variation across speech and writing.* Cambridge: Cambridge University Press.

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

———. 2003. Variation among university spoken and written registers: A new multi-dimensional analysis. In *Corpus analysis: Language structure and language use*, ed. C. Meyer and P. Leistyna, 47–70. Amsterdam: Rodopi.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, and Edward Finegan. 1999. *Longman grammar of spoken and written English*. London: Longman.

Biber, Douglas, Susan Conrad, Randi Reppen, Patricia Byrd, and Marie Helt. 2002. Speaking and writing in the university: A multidimensional comparison. *TESOL Quarterly* 36:9–48.

Biber, Douglas, Eniko Csomay, James K. Jones, and Casey Keck. 2004. Vocabulary-based discourse patterns in university registers. In *Corpora and discourse,* ed. Allen Partington, John Morley, and L. Haarman. Bern: Peter Lang.

Csomay, Eniko. 2000. Academic lectures: An interface of an oral/literate continuum. *NovELTy,* 7, no. 3:30–48.

———. 2002a. Variation in academic lectures: interactivity and level of instruction. In *Using corpora to explore linguistic variation*, ed. Randi Reppen, Susan Fitzmaurice and Douglas Biber, 203–24. Philadelphia: John Benjamins.

———. 2002b. Episodes in university classrooms: A corpus-linguistic investigation. Ph.D. diss., Northern Arizona University.

DeCarrico, Jeanette, and James R. Nattinger. 1988. Lexical phrases for the comprehension of academic lectures. *English for Specific Purposes* 7:91–102.

Fox, Barbara A. 1987. *Discourse structure and anaphora. Written and conversational English.* Cambridge: Cambridge University Press.

Hansen, Christa. 1994. Topic identification in lecture discourse. In *Academic listening: research perspectives,* ed. John Flowerdew. New York: Cambridge University Press.

Hearst, Marti A. 1997. TextTiling: Segmenting text into multi-paragraph subtopic passages. *Computational Linguistics* 23, no. 1:33–64.

Mann, William C., Christian M. Matthiessen, and Sandra A. Thompson. 1992. Rhetorical structure theory and text analysis. In *Discourse description: Diverse linguistic analyses of a fund-raising text,* ed. William C. Mann and Sandra A. Thompson. Amsterdam: John Benjamins.

Poole, Deborah. 2003. Linguistic connections between co-occurring speech and writing in a classroom literacy event. *Discourse Processes* 35, no. 2: 103–4.

Reppen, Randi, Susan Fitzmaurice, and Douglas Biber, eds. 2002. *Using corpora to explore linguistic variation.* Philadelphia: John Benjamins.

Tyler, Andrea. 1995. Patterns of lexis: How much can repetition tell us about discourse coherence? In *Linguistics and the education of language teachers: Ethnolinguistic, psycholinguistic, and sociolinguistic aspects,* ed. J. E. Alatis, C. A. Straehle, B. Gallenberger, and M. Ronkin. Georgetown University Round Table on Languages and Linguistics. Washington, D.C.: Georgetown University Press.

Youmans, Gilbert. 1991. A new tool for discourse analysis: The Vocabulary Management Profile. *Language* 67:763–89.

Young, Lynne. 1994. University lectures—macro-structure and micro-features. In *Academic listening: Research perspectives*, ed. John Flowerdew, 159–76. New York: Cambridge University Press.