SSLA, **24**, 199–208
DOI: 10.1017.S0272263102002048

WHAT DOES FREQUENCY HAVE TO DO WITH GRAMMAR TEACHING?

Douglas Biber and Randi Reppen
Northern Arizona University

Using frequency findings from corpus linguistics, this paper explores the relationship between the information presented in ESL-EFL materials and what is known about actual language use based on empirical studies. Three aspects of materials development for grammar instruction are discussed: the grammatical features to be included, the order of grammatical topics, and the vocabulary used to illustrate these topics. For each aspect, we show that there are often sharp contrasts between the information found in grammar materials and what learners encounter in the real world of language use. In our conclusion, we argue that a selective revision of pedagogy to reflect actual use, as shown by frequency studies, could result in radical changes that facilitate the learning process for students.

The development of materials for language instruction and assessment requires materials writers and teachers to make repeated judgments about language use and to make decisions about the linguistic features and words to include. In recent years, most ESL professionals have adopted a preference for authentic materials, presenting language from natural texts rather than made-up examples (Byrd, 1995b; McDonough & Shaw, 1993). However, there are additional underlying decisions required concerning the language forms to be presented at any given stage. For example, authors of grammar textbooks regularly confront three basic issues:

1. Which grammatical features should be included in a lesson or book; which should be excluded? How much space should be given to included features?
2. What should the order of grammatical topics be?
3. Which specific words should be included when illustrating a grammatical feature?

Address correspondence to: Douglas Biber, English Department, Box 6032, Northern Arizona University, Flagstaff, AZ 86011–6032; e-mail: douglas.biber@nau.edu.

© 2002 Cambridge University Press 0272-2631/02 $9.50

Textbooks for materials writers provide little help for authors making these decisions. In fact, most treatments fail to even acknowledge these issues. Dubin (1995) noted that textbook authors are "on their own, so to speak, when it comes to making decisions about a number of issues. What they lack is access to a well-developed body of knowledge about materials writing" (p. 15). As a result, these decisions have usually been based on the author's gut-level impressions and anecdotal evidence of how speakers and writers use language. These impressions usually operate below the level of consciousness and are often regarded as accepted truths. As Byrd (1995a) wrote, "often design decisions are based on traditions about grammar materials and their organization rather than on careful rethinking of either the content or its organization" (p. 46).

One empirical basis that could be used for many of these decisions is frequency information. Over the past 10–20 years, empirical analyses of representative corpora have provided a wealth of information about the actual patterns of language use in English. Biber, Conrad, and Reppen (1998) and Kennedy (1998) provide useful introductions to this analytical approach, whereas the edited volumes by Simpson and Swales (2001), Aarts and Meyer (1995), Aijmer and Altenberg (1991), Johansson and Stenström (1991), and Granger (1998) provide good collections of research studies of this type. There are also a number of book-length treatments reporting corpus-based investigations of grammar and discourse (see, e.g., Tottie, 1991, on negation; Collins, 1991, on clefts; Granger, 1983, on passives; Mair, 1990, on infinitival complement clauses; Meyer, 1992, on apposition; Partington, 1998, on collocational patterns). Multidimensional analysis was developed as a quantitative, corpus-based methodology to study the coordinated patterns of use among a full range of linguistic features. Multidimensional studies have documented the underlying patterns of variation among spoken and written registers (e.g., Biber, 1988, 1995), and they have also been used to investigate more specialized issues relating to ESP-EAP, language development, and language testing (see Conrad & Biber, 2001). Finally, the recent *Longman Grammar of Spoken and Written English* (*LGSWE*) (Biber, Johansson, Leech, Conrad, & Finegan, 1999) applies corpus-based analysis to grammatical description, showing how any grammatical feature can be described for both structural characteristics and discourse patterns of use.

However, as Ellis (pp. 175–178) points out, frequency information has been disregarded in applied linguistics over the past two decades, largely because of associations with the Audio-Lingual Method (Lado, 1964) and behaviorism more generally (Skinner, 1957). This disregard is often accompanied by the assumption that current pedagogical practice is more carefully grounded than a reliance on frequency. In practice, though, this is not the case. Rather, authors often make pedagogical decisions based on their beliefs about language use, in many cases without even acknowledging that decisions are being made. Unfortunately, our intuitions about language use are often wrong. As a result, teaching and assessment materials often fail to provide an accurate reflection of the language actually used by speakers and writers in natural situations.

The present paper illustrates the kinds of unexpected findings about language use that result from corpus-based investigations and how such findings lead to fundamentally different decisions about materials design. Specifically, we contrast the presentation of information in six ESL grammar textbooks with empirical frequency findings based on corpus research done for the *LGSWE* (Biber et al., 1999). Three case studies are considered, one for each of the major issues identified earlier: grammatical features to include or exclude (focusing on noun premodifiers); the order of grammatical topics (focusing on progressive and simple present tense); and specific words to include when illustrating a grammatical feature (focusing on the verbs used in the discussion of present progressive and simple present tense). For all three issues, we show how corpus-based frequency findings lead to decisions radically different from the current practice of existing materials. That is, although frequency information can never be the sole factor used to design materials, it does provide a more solid basis than relying only on intuitions and accepted practice.

METHODOLOGY

To get an idea of current practice in grammar instruction, we surveyed six popular ESL-EFL grammar books: *Basic Grammar in Use* (Murphy, Altman, & Rutherford, 1989) for low intermediate; *Focus on Grammar* (Fuchs, Bonner, & Westheimer, 1999), *Fundamentals of English Grammar (*Azar, 1992), and *English Grammar in Use* (Murphy, 1986) for intermediate; *Grammar Dimensions 3* (Thewlis, 2000) and *Oxford Practice Grammar* (Eastwood, 1992) for intermediate to advanced.

This sample includes texts that are widely used and highly regarded as clear, effective treatments. (We avoided high advanced books, which focus on more specialized topics.) Even though this is clearly not an exhaustive sample, it allows us to characterize current practice to the extent that consistent approaches are followed across these texts.

The priorities of these ESL textbooks are compared to the frequency findings reported by Biber et al. (1999). The *LGSWE* reports the results of corpus-based analyses based on approximately 20 million words from four registers: conversation, fiction, newspaper language, and academic prose (see Biber et al., chap. 2, for a complete description of the corpus). All frequency counts reported here have been normalized to a common basis (a count per 1 million words of text) so that they are directly comparable across registers.

RESULTS

Which Grammatical Features Should Be Included and Which Should Be Excluded in a Lesson or Book?

Adjectives are typically characterized as words that describe something, and attributive adjectives are presented as the major grammatical device used for

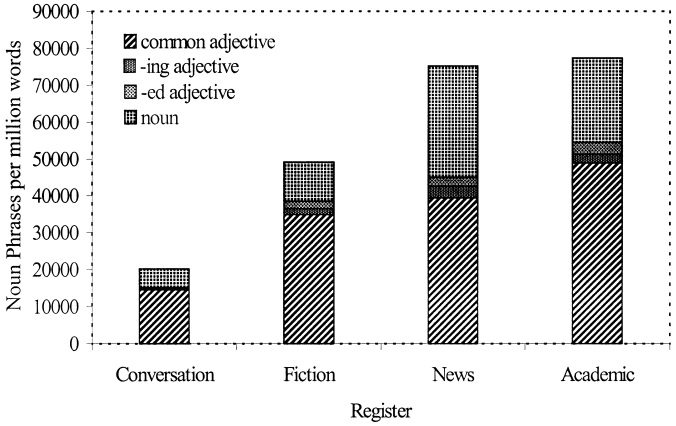

Figure 1. Frequencies of adjectives, participial adjectives, *-ed* adjectives, and nouns as nominal premodifiers (based on Biber et al., 1999, Figure 8.7).

noun modification (e.g., *the big house*). Most textbooks also describe the adjectival role of *-ing* and *-ed* participles (e.g., *an exciting game*, *an interested couple*). Thus, of the six textbooks that we surveyed, five included sections on adjectives and four included discussion of participial adjectives. The adjectival role of nouns (e.g., *a grammar lesson*) is less commonly acknowledged in textbooks; only one of the textbooks in our survey included discussion or illustration of nouns functioning as nominal premodifiers. This difference seems to reflect a widely held belief that adjectives and participial adjectives are the primary devices used for noun modification, whereas nouns are considered to be much less important in this role.

However, corpus-based analysis provides a very different picture. Figure 1 presents the frequencies of adjectives, participial adjectives, and nouns as nominal premodifiers. In conversation, adjectives are the primary device used for noun modification (although most noun phrases in conversation do not include modifiers). Given that conversational English is the primary target for lower level textbooks, an exclusive focus on adjectives seems justified in those books. A dramatically different pattern of use is found in the written registers, however. The pattern in newspaper writing is especially noteworthy: Nouns as premodifiers are extremely frequent and nearly as common as adjectives, whereas participial forms are surprisingly rare.

It might be argued that the grammar of nouns as premodifiers is somehow simpler than that of adjectives or participial forms, and therefore they require little overt attention. However, in actual fact, premodifying nouns can express a bewildering array of meanings, with no surface-level clues to guide the reader. For example, consider the relationships between the modifying noun (N1) and the head noun (N2) in the pairs in (1).

(1)	*glass windows*, *metal seat*, *tomato sauce*	(N2 is made from N1)
	pencil case, *brandy bottle*, *patrol car*	(N2 is used for the purpose of N1)
	sex magazine, *sports diary*	(N2 is about N1)
	farmyard manure, *computer printout*	(N2 comes from N1)
	summer rains, *Paris conference*	(N1 gives the time or location of N2)

These are only a few of the many different meaning relations found with nouns as premodifiers (see Biber et al., 1999, pp. 589–591). Thus, these forms are potentially difficult to understand in addition to being extremely frequent in written registers. It seems obvious that students at intermediate and advanced levels need greater exposure to these commonly encountered forms than comparatively rare forms like participial adjectives.

What Should the Order of Grammatical Topics Be?

One of the most widely held intuitions about language use among TESL professionals is the belief that progressive aspect is the unmarked choice in conversation. This belief is sometimes reflected in the extremely frequent use of progressive verbs (underlined in example [2]) in made-up dialogues like those found in ESL-EFL course books teaching conversation skills. For example, consider the conversation in (2), from *As I Was Saying: Conversation Tactics* (Richards & Hull, 1987):

(2) Doctor: *Hello Mrs. Thomas. What can I do for you?*
Patient: *Well, I've <u>been having</u> bad stomach pains lately, doctor.*
Doctor: *Oh I'm sorry to hear that. How long have you <u>been having</u> them?*
Patient: *Just in the last few weeks. I get a very sharp pain about an hour after I've eaten.*
[. . .]
Doctor: *Well, I don't think it's anything serious. Maybe you eat too quickly. You don't give yourself time to digest your food.*
Patient: *My husband <u>is always telling</u> me that.*

This belief is similarly reflected in the sequence of topics found in most ESL grammar books, in which the progressive is presented as one of the fundamental building blocks of English grammar. For example, four of the six ESL textbooks in our survey introduce the progressive in the first chapter. Three of these books introduce the progressive before covering the simple present; two others introduce the progressive in the same chapter as the simple present. Given the nature of this coverage, it would be entirely natural for learners to use progressive verbs as their first choice, at least in conversation.

As Figure 2 shows, the generalization that progressive aspect is more common in conversation than in other registers is correct. The contrast with academic prose is especially noteworthy: Progressive aspect is rare in academic prose but common in conversation. However, as Figure 3 shows, it is not at all correct to conclude that progressive aspect is the unmarked choice in conversation. Rather, simple aspect is clearly the unmarked choice. In fact, simple

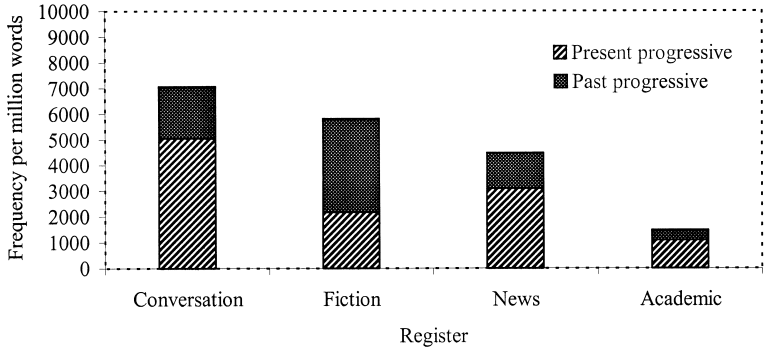

Figure 2. Frequency of past and present progressive aspect across registers (based on Biber et al., 1999, Figure 6.4).

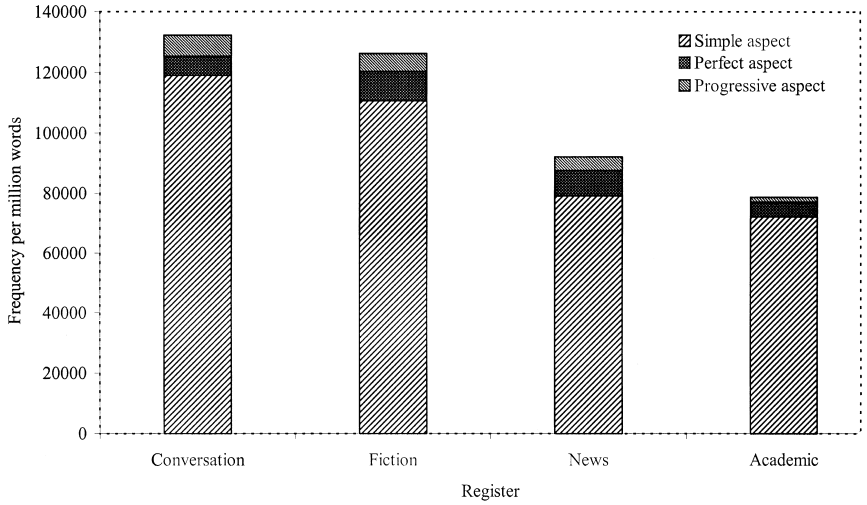

Figure 3. Frequency of simple, perfect, and progressive aspect in four registers (based on Biber et al., 1999, Figure 6.2).

aspect verb phrases are more than 20 times as common as progressives in conversation. The excerpt in (3) illustrates this extreme reliance on simple aspect in natural conversation.

(3) B: *What do you do at Dudley Allen then?*
A: *What the school?*
B: *Yeah. Do you—*
A: *No I'm, I'm only on the PTA.*
B: *You're just on the PTA?*

A: *That's it.*
B: *You don't actually work?*
A: *I work at the erm—*
B: *I know you work at Crown Hills, don't you?*
A: *Yeah.*

The prominence given to progressive aspect in textbooks seems to be based on the mistaken belief that it is fundamentally important to communication, at least in conversation. At the same time, it is not at all uncommon to hear teachers commenting on the overuse of the progressive by students. We would argue that such overuse is not surprising, given that instructional materials implicitly suggest that progressives are more important (and far more common) than they actually are.

Which Specific Words to Include When Illustrating a Grammatical Feature?

There are literally hundreds of common lexical verbs in English. For example, nearly 400 different verb forms occur over 20 times per million words in the *LGSWE* corpus (see Biber et al., 1999, pp. 370–371). These include many everyday verbs, such as *pull*, *throw*, *choose*, *fall*, and so forth. Given this large inventory of relatively common verbs, it might be easy to assume that no individual verbs stand out as being particularly frequent. However, this is not at all the case: There are only 63 lexical verbs that occur more than 500 times per million words in a register, and only 12 lexical verbs occur more than 1,000 times per million words in the *LGSWE* corpus (Biber et al., pp. 367–378). These 12 most common lexical verbs are: *say*, *get*, *go*, *know*, *think*, *see*, *make*, *come*, *take*, *want*, *give*, and *mean*. (Additionally, the primary verbs *be*, *have*, and *do* are extremely common.)

To give an indication of the importance of these 12 verbs, Figure 4 plots their combined frequency compared to the overall frequency of all other verbs. Taken as a group, these 12 verbs are especially important in conversation, where they account for almost 45% of the occurrences of all lexical verbs. Obviously, any conversational primer that did not include extensive practice of these words would be shortchanging students.

An important function of grammar textbooks is the introduction of new vocabulary. We would argue that frequency is an important guiding principle here: Frequent words will be more useful to students receptively and in production, whereas relatively rare words will prove less useful in the earlier stages of language learning. However, a survey of the six low-intermediate and intermediate books in our sample suggests that there is little consistency across books guiding the selection of illustrative vocabulary. We focused specifically on the verbs illustrated in the initial lessons describing present progressive and simple present tense. Some of the most common verbs were included in at least some of these textbooks, including: *take*, *come*, *like*, *want*, *know*, and *mean* (and the primary verbs *be*, *have*, and *do*). However, 7 of the

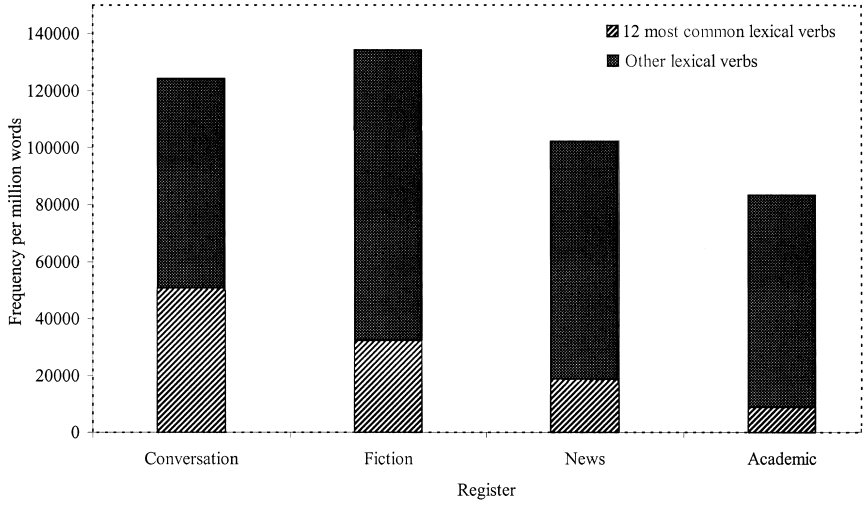

Figure 4. Proportion of most common lexical verbs across four registers (based on Biber et al., 1999, Figure 5.8)

12 most common lexical verbs were disregarded by all textbooks in our survey: *get*, *go*, *see*, *make*, *give*, *say*, and *think*. Additionally, many moderately common verbs (e.g., *try*, *put*, *use*, *leave*) are disregarded in most books. In contrast, many textbooks include examples containing relatively rare verbs, such as *wear*, *cry*, *revolve*, *arrive*, *touch*, *travel*, *read*, *rain*, *shine*, *write*, *ring*, *drive*, *enjoy*, *study*, *build*, *rise*, *smoke*, *close*, *speak*, *grow*, *kiss*, *stay*, *own*, *taste*, *cause*, and *boil*. Although some of these verbs seem common to native speakers, they in fact occur much less frequently. (Note that there are no major grammatical differences between the included and excluded sets of verbs: Both sets include regular and irregular verbs, transitive and intransitive verbs, and verbs that take complex complementation patterns.) We of course would not argue against inclusion of a wide range of vocabulary, given students' fundamental need to acquire vocabulary knowledge. However, we can imagine no reason why relatively rare words should be illustrated to the exclusion of common words in lower level books. Here again, we see how authors have been forced to rely on intuitions regarding the typical patterns of language use, because they have lacked frequency information about the most common forms with a given grammatical pattern.

CONCLUSION

Ellis has shown us that language learners naturally rely on frequency for many different language tasks, ranging from irregular verb patterns to collocation patterns. Given its importance in acquisition, we would argue that frequency

should also play a key role in the development of materials and in the choices that teachers make in language classrooms. With the recent availability of comprehensive frequency-based grammatical descriptions, such integration of pedagogy and research has become feasible.

Actually, textbook authors and students have been interested in frequency information all along. For example, many books include lists of common phrasal verbs or provide other information about patterns that are common or typical. However, more often than not, this information is based on the author's intuitions rather than empirical research. As a result, these lists often include forms that are not in fact common, while overlooking other forms that do occur frequently. (Surprisingly, extremely frequent forms are also the ones that we are most likely to overlook.)

By using information based on actual frequency and context of use (e.g., register differences), materials developers and teachers should be able to increase the meaningful input that is provided to learners. Obviously, other factors are equally important—for example, some grammatical topics are required as building blocks for later topics; some grammatical topics are more difficult and therefore require more practice than others. In many cases, though, language use has been a primary guiding principle: Authors have attempted to present the typical and most important patterns first, moving on to more specialized topics in later chapters and more advanced books. Lacking empirical studies, authors have been forced to rely on their intuitions for these judgments about language use, and widely accepted norms have arisen to support those intuitions. With the rise of corpus-based analysis, we are beginning to see empirical descriptions of language use, identifying the patterns that are actually frequent (or not) and documenting the differential reliance on specific forms and words in different registers. In some cases, our intuitions as authors have turned out to be correct; in many other cases, we have been wrong. For those latter cases, revising pedagogy to reflect actual use, as shown by frequency studies, can result in radical changes that facilitate the learning process for students.

REFERENCES

Aarts, B., & Meyer, C. (Eds.). (1995). *The verb in contemporary English: Theory and description*. New York: Cambridge University Press.

Aijmer, K., & Altenberg, B. (Eds.). (1991). *English corpus linguistics*. London: Longman.

Azar, B. (1992). *Fundamentals of English grammar*. Englewood Cliffs, NJ: Prentice-Hall.

Biber, D. (1988). *Variation across speech and writing*. New York: Cambridge University Press.

Biber, D. (1995). *Dimensions of register variation: A cross-linguistic perspective*. New York: Cambridge University Press.

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. New York: Cambridge University Press.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. London: Longman.

Byrd, P. (1995a). Issues in the writing of grammar textbooks. In P. Byrd (Ed.), *Materials writer's guide* (pp. 45–63). Boston: Heinle & Heinle.

Byrd, P. (Ed.). (1995b). *Materials writer's guide*. Boston: Heinle & Heinle.

Collins, P. (1991). *Cleft and pseudo-cleft constructions in English*. London: Routledge.

Conrad, S., & Biber, D. (Eds.). (2001). *Multi-dimensional studies of register variation in English*. London: Longman.
Dubin, F. (1995). The craft of materials writing. In P. Byrd (Ed.), *Materials writer's guide* (pp. 13–22). Boston: Heinle & Heinle.
Eastwood, J. (1992). *Oxford practice grammar*. Oxford: Oxford University Press.
Fuchs, M., Bonner, M., & Westheimer, M. (1999). *Focus on grammar*. London: Longman.
Granger, S. (1983). *The BE+past participle construction in spoken English with special emphasis on the passive*. Amsterdam: Elsevier Science.
Granger, S. (Ed.). (1998). *Learner English on computer*. London: Longman.
Johansson, S., & Stenström, A.-B. (Eds.). (1991). *English computer corpora: Selected papers and research guide*. Berlin: Mouton de Gruyter.
Kennedy, G. (1998). *An introduction to corpus linguistics*. London: Longman.
Lado, R. (1964). *Language teaching: A scientific approach*. Boston: McGraw-Hill.
Mair, C. (1990). *Infinitival complement clauses in English*. New York: Cambridge University Press.
McDonough, J., & Shaw, C. (1993). *Materials and methods in ELT*. Oxford: Blackwell.
Meyer, C. (1992). *Apposition in contemporary English*. New York: Cambridge University Press.
Murphy, R. (1986). *English grammar in use*. New York: Cambridge University Press.
Murphy, R., Altman, R., & Rutherford, W. (1989). *Basic grammar in use*. New York: Cambridge University Press.
Partington, A. (1998). *Patterns and meanings: Using corpora for English language research and teaching*. Amsterdam: Benjamins.
Richards, J., & Hull, J. (1987). *As I was saying: Conversation tactics*. Reading, MA: Addison-Wesley.
Simpson, R., & Swales, J. (Eds.). (2001). *Corpus linguistics in North America*. Ann Arbor: University of Michigan Press.
Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.
Thewlis, S. (2000). *Grammar dimensions 3*. Boston: Heinle & Heinle.
Tottie, G. (1991). *Negation in English speech and writing: A study in variation*. San Diego, CA: Academic Press.