Simplicity Without Elegance: Features of Sentences in L1 and L2 Academic Texts

ELI HINKEL
Seattle University
Seattle, Washington, United States

A quantitative analysis of 1,083 L1 and L2 academic texts establishes that advanced nonnative-English-speaking students in U.S. universities employ excessively simple syntactic and lexical constructions, such as *be*-copula as the main verb; predicative adjectives; vague nouns; and public, private, and expecting/tentative verbs, at median frequency rates significantly higher than those found in basic texts by native English speakers. An examination of substantial corpus analyses carried out in the past two decades indicates that these constructions are prevalent in conversational and informal discourse rather than written academic texts. Reasons for the prevalence of simple syntactic and lexical features of text in L2 academic essays are considered. In addition, instructional techniques are proposed to deal with shortfalls in naturalistic and communicative L2 learning and instructional methods for academically bound L2 students.

In his overview of research into L2 academic writing, Silva (1993) points out, "There also seems to be a clear need for more extensive treatment of textual concerns. . . . It may be also necessary for L2 writing teachers to work to enhance their writers' grammatical and lexical resources" (p. 671) to allow students to build a syntactic and lexical repertoire with which to produce more sophisticated academic texts. One of the difficult issues in teaching academically bound ESL students to produce appropriate academic written text is that research has not established with certainty what specific syntactic and lexical features, when taken together, can create an impression of a seemingly simplistic or reasonably sophisticated text in written L2 discourse. This article, therefore, identifies the grammar and vocabulary constructions that may create an overall impression of textual simplicity in texts written by nonnative English speakers (NNSs) relative to those written by native speakers (NSs) and that may therefore reflect negatively on the quality of

L2 academic essay text. Results from a quantitative analysis of frequency rates of syntactically and lexically simple constructions in a corpus of L1 and L2 academic texts suggest several specific and clear-cut syntactic and lexical text features that may be explicitly targeted in instruction to help NNS writers improve the overall quality of their text.

THE SIGNIFICANCE OF COMPLEXITY IN WRITING

Investigations into L2 writing and text have established that in large-scale testing and university-level assessments of student essays, syntactic and lexical simplicity is often considered to be a severe handicap, and research has shown that essay raters almost always note simple constructions and lexicon, a consideration that may reduce the rating (Reid, 1993; Vaughan, 1991). Several Test of English as a Foreign Language (TOEFL) research publications aimed at better understanding variation in writing quality identify many characteristics of simple or sophisticated uses of language. In these studies, the frequency rates of rare and derived words ("unique words" and "longer words") are considered to be markers of L2 writers' broad vocabulary ranges and stylistically "'precise' expressions" have been shown to be reliable predictors of overall Test of Written English scores (Frase, Faletti, Ginther, & Grant, 1999, pp. 20–22). These and other variables, such as accurate and extensive use of subordinate clauses and appropriate use of articles, contribute to a higher degree of text sophistication. In general, the TOEFL evaluative criteria developed for ESL writing and language use include such parameters as "threshold levels of vocabulary established in relation to word frequency counts of corpora of English texts or to readers' impressions of the vocabulary" but "not overuse" of more frequent items, such as *know, say,* and *think* (Cumming, Kantor, Powers, Santos, & Taylor, 2000, p. 15). Hamp-Lyons (1991) and Davidson (1991) refer to important disadvantages associated with use of simple constructions and vocabulary in essay assessments. According to Read (2000), in standardized tests of ESL writing, a great deal of importance is attached to "lexical richness" (p. 200), which is defined as a relatively high percentage of low-frequency words appropriate to the topic and style as opposed to a preponderance of high-frequency, everyday words.

In the teaching of L2 writing, the issues of fluency, syntactic and lexical accuracy, variety, and sophistication play a crucial role (Ferris & Hedgcock, 1998; Jordan, 1997; Reid, 1993). A number of researchers have found that academic texts written by NNSs frequently rely on a limited lexical repertoire that results in vague and less sophisticated prose relative to that of NSs (Carlson, 1988; Leki, 1991; Read, 2000). Similarly, Johnson and Roen (1989) point out that instructors of English

mainstream and discipline-specific courses often see the written discourse of NNS students as basic, limited, and unskilled, and view many L2 writers as being unable to produce effective text.

For developing effective instructional models, Pienemann (1985) points out that syntactic and lexical simplicity criteria in formal language learning serve as "less ambiguous principles" (p. 24) for determining L2 syllabus and instructional foci. He notes that by means of an analysis of the simple linguistic features that occur in L2 production, curriculum and syllabus development can focus on those facets of language that can benefit from explicit teaching, such as L2 syntactic regularities and expanded lexical repertoire. Specifically, he calls for instruction to expand the range of features not found in naturally occurring (and predominantly spoken) discourse and comments that an identification of simple and frequent grammatical and lexical points can serve as the basis for the teaching of more advanced linguistic features, thereby improving the quality of L2 instruction and learning.

In empirical investigations, Pica's (1985) study shows that grammatical simplicity and a preponderance of simple structures in L2 production represent an important measure for identifying L2 features that academically and professionally bound students need to learn. In her view, as in Pienemann's (1985), an examination of syntactically simple features that occur in L2 language uses can contribute to developing L2 syllabuses with explicit and grounded foci.

The features of syntactic and lexical accuracy and fluency serve as foundational constructs for a model designed by Wolfe-Quintero, Inagaki, and Kim (1998) to measure developmental proficiency, fluency, and accuracy as well as "basicness" (p. 106) of L2 written text. Although the enormous complexity of this multifaceted model crucially hinges on the interactions among these important constructs, one of the model's interesting features is that it includes measures of lexical variation, sophistication, and basicness that are related to L2 proficiency. For example, to measure lexical variation, the model takes into account the number of times a particular lexical item is repeated[1] per number of words in a written text, and lexical sophistication and basicness are measured by means of a complex ratio of the number of low- and high-frequency words encountered per total number of words in a writing

---

[1] Reynolds (2001), however, showed that counter to Wolfe-Quintero et al.'s (1998) model, word and lexical repetition measures are not necessarily accurate or consistent predictors of ESL writing development or basicness. His type/token analysis of lexical repetition in 191 NS and NNS essays shows that in ESL texts, lexical item repetition occurs in T-unit clusters and depends on the essay topic and rhetorical pattern, as well as writers' cultural backgrounds. Specifically, in NNS texts, increased T-unit length (but not text length) determines significantly higher repetition values. The discourse functions of repetition in NNS students' writing are discussed in Reynolds (1995).

sample. However, the authors note that measures of variation and sophistication often result in particularly low rates of occurrences of sophisticated lexical items in L2 written text (in most cases, substantially fewer than 1 word per 100). An important problem encountered by researchers is that, to be valid, measures of lexical variation and sophistication should be applied to texts of at least 350 words in length; in Wolfe-Quintero et al.'s model, most L2 written texts were markedly shorter.

IDENTIFYING SIMPLICITY IN TEXT

Since at least the 1920s, a great deal of effort has been devoted to identifying syntactically and lexically simple features of text to use in developing L2 material and teaching that would reduce the learning burden or simplify communication in an L2. Researchers have persistently investigated a connection between word frequencies and the simplicity of their lexical content since the 1930s, when the pedagogical and communicative purpose of identifying the most common and simple words in English was to simplify texts to accommodate learners with limited vocabulary ranges. One of the most prominent first attempts at simplified word lists was Ogden's *Basic English* (1930), which had the goal of compiling a list of 850 simple and essential words for communicating with "foreigners": "English-speaking people would do well to keep their own English as near as possible to the simplicity and precision of Basic English" (p. 6). Ogden's list was followed by another attempt to develop a list of vocabulary items for producing simple reading materials (Palmer, West, & Faucett, 1936) and later by West's (1953) General Service List, which used a far more systematic approach to the analysis of word frequencies to develop a list of the most useful and common vocabulary for L2 learning. West's list centered on the semantic count of words by their meanings and had the stated goal of eliminating "less frequent" and "heavy words" to facilitate vocabulary learning by enhancing the "ease" and "clearness" of word meanings (pp. viii–ix).

In the 1980s, pedagogical materials for L2 learning identified lists of basic nouns and verbs as those that referred to frequently encountered and concrete everyday objects, actions, and ideas to meet the needs of beginning learners for "simple communication experiences" (Allen, 1983, p. 18) in conversations with NSs. On the other hand, in recent years, the research in ESL pedagogy on identifying simple and complex grammatical structures and vocabulary has been motivated by the converse goal of helping learners improve the quality and sophistication of their language production and text (Hamp-Lyons, 1991; Jordan, 1997; Nation, 1990, 2001).

To date, numerous competing definitions of simple (or complex) syntactic constructions and words have been developed, and to account for all (or even many) in any degree of depth here would be practically impossible. For example, in L2 material development, teaching, and corpus analyses, some definitions use frequencies of syntactic structures and words to identify the most common, useful, and practical ones for learners (Nation, 1990; Thorndike & Lorge, 1944; West, 1953); others focus on the grammar and lexis minimally essential to express ideas (Allen, 1983; Ogden, 1936; Richards, 1943); and still others deal with the simplicity (or complexity) of syntactic and morphological derivations (Adams, 2001; Biber, Johansson, Leech, Conrad, & Finegan, 1999; Quirk, Greenbaum, Leech, & Svartvik, 1985). In analyses of ESL essays and teaching, all such definitions have advantages and shortcomings because none is based on actual spoken or written language produced by learners in a specific context but rather on language produced by NSs.

The features selected for inclusion in this analysis of NS and NNS student essays are based on a few specific, narrowly defined types of syntactic structures and lexical classes of words that have been identified as comparatively simple and common in large corpus analyses of spoken and written English published in the past decade. In particular, the study does not focus on individual sentences or words but on entire syntactic and lexical classes of high-frequency features prevalent in conversational and spoken registers and popular print media (see the Methods section). Like other attempts to define simple or common features identified in various corpora, this approach has advantages and shortcomings: On the one hand, the inclusion of feature classes permits the making of some generalizations, and, on the other, individual sentences or words in any class of items can have higher frequencies of use than others.

For NNS students in U.S. universities, exposure to and experience with L2 spoken and formal written registers may represent a key consideration that can ultimately determine the types of syntactic and lexical features such students employ in L2 essays. In a recent study, Shaw and Liu (1998) examined issues of textual simplicity and complexity in academic essays written by NNS university students. Shaw and Liu's study, based on pre- and posttest essays written by 164 speakers of 23 languages, found that the students' uses of informal linguistic features actually increased in frequency with greater exposure to the spoken register in English and despite intensive academic writing instruction. On the other hand, the researchers did not identify a similar increase in the use of textual features of "academic style" (p. 246), which are considered to be lexically and syntactically complex (e.g., subordinate clauses), whereas the rates of occurrence of spoken and informal syntactic and lexical features (i.e., informal vocabulary items, such as *a lot, because of,* and *so*) increased significantly. The study concludes that in English for academic

purposes programs in English-speaking countries, students are exposed to formal English writing but also have much contact with informal conversational discourse. The authors point out that L2 learners become adept at employing the features of the conversational register without developing register differentiation skills.

In the past two decades, corpus analyses of spoken and written text in English have made great strides in identifying the contextual meanings and functions of syntactic and lexical features that are prevalent in the conversational genre and that can make written academic text appear to be particularly simple. According to Biber et al. (1999), for example, compared with frequencies of lexical items in any other register (e.g., fiction or news), the conversational register shows a clear preference for simple nouns and noun phrases, and derived nouns are by far the least common in conversations.

In investigations of features in a corpora of social conversations and spoken discourse, Brazil (1995) notes the preponderance of existential constructions, such as *be*-copula + predicative adjective structures and *there*-existential slot filler + *be*-copula, as well as other types of *be*-copula constructions with adverbs and prepositional phrases. According to Brazil, the function of the *be*-copula structures is to specify identities, qualities, conditions, and locations that are often indeterminate and serve as a limited means of providing static descriptions of states and situations. Brazil also comments on the frequency of "things," "people," and "places" nouns that are assigned "substantive status," or "everything that is traditionally thought of as . . . a noun" (p. 151). In Sinclair's (1991) analyses of large spoken and written corpora, the list of the most frequent 113 forms that appear among approximately 30 million contains such nouns as *human, man,* and *people* and such verbs as *be, know, say, see,* and *think* (in all forms).

The most extensive study of significant differences between the frequencies of lexical items in speech and writing was carried out by Leech, Rayson, and Wilson (2001). They found that finite forms of *be* (e.g., *am, are, is, was, were*) were far more common in spoken than in written discourse, as were, for example, other highly frequent items, such as various forms of *agree, ask, believe, boy, call, decide, difference, expect, feel, fact, forget, girl, hear, hope, idea, know, learn, like, listen, look, love, man, mention, people, plan, question, read, remember, say, see, speak, study, stuff, suppose, talk, tell, thing, think, try, want, way, whatever, wish, woman, wonder, word,* and *write.* (All of these items, together with other members of their broad lexical classes, were frequently encountered in the NS and NNS essays analyzed in this study.)

METHOD

The study compared the median frequency rates of simple syntactic and lexical features employed in academic essays of students who were speakers of six languages: NSs of American English, Chinese, Japanese, Korean, Indonesian, and Arabic. The analysis identified specific features of L2 grammar and vocabulary that can be useful in L2 instruction and can simultaneously meet several L2 pedagogical goals: bring learners' attention to issues of divergent L2 registers and genres, focus on syntactic and lexical manifestations of various registers and genres in text, and emphasize the importance of appropriate grammar and lexical range in written academic text.

Syntactic and Lexical Forms Investigated

Based on the findings of various analyses of spoken and written corpora, the study focused on the frequency rates of *be*-copula as a sentence main verb and its attendant structures, that is, predicative adjectives and existential *there*. However, *be*-copula also occurs in the advanced and sophisticated construction *it*-cleft, the uses of which need to be accounted for with the purpose of determining whether its presence in L1 and L2 academic essay texts affects the frequencies of *be*-copula use. In addition, the discussion below addresses the frequency rates of nouns and verbs in high-frequency lexical classes: vague nouns (*people, thing, way*) (Brazil, 1995; Channell, 1994; Francis, 1994; Read, 2000; Sinclair, 1991), public verbs (*say, state, talk*) (Hunston & Francis, 1996; Quirk et al., 1985), private verbs (*feel, learn, study*) (Leech et al., 2001; Quirk et al., 1985), and expecting/tentative verbs (*like, try, want*) (Biber et al., 1999; Leech et al., 2001).

Simple Syntactic Features

Constructions with *be* as a main verb can have copula or existential functions and usually mark a static informational style in written prose (Quirk et al., 1985). An example (from a text written by a Korean NS) is *In Korea, some families are very rich, and their children are lazy because they don't have to do anything.* Biber (1988) calls these "non-complex constructions" (p. 228) because they have reduced information content and are characteristic of spoken discourse. In descriptive and expository prose, such constructions have existential meanings when they supply or present information. According to Biber, structures with *be* as the main verb are somewhat simpler than those with verbs that have a higher

semantic and lexical content, particularly so in descriptive contexts with predicative adjectives (e.g., *children are happy—happy children*). In this study, the counts of *be*-copula included all forms of *be* as the main verb in all tenses (and in contractions).

Predicative adjectives are those that form part of a clause predicate, following *be*-copula or linking verbs (e.g., *become, grow, seem/appear*); for example, *Students are responsible for their studying, grades or not* (from a text written by an Arabic NS). Corpus analysis findings (Biber et al., 1999) have determined that the *be*-copula "is overwhelmingly the most common verb" (p. 437) and that it takes predicative adjectival complements 20 times more often than other copular verbs. Predicative adjectives are distributed approximately evenly across conversational and academic genres. However, most predicative adjectives are found in contexts that describe a state of mind or emotion (e.g., *afraid, anxious, happy, sad, sure*) that is common in spoken and conversational registers (Biber et al., 1999; Quirk et al., 1985) and infrequently expressed with depersonalized *it*-cleft structures (see below). The use of predicative adjectives limits the range and type of content that can be conveyed because they require the presence of copular verbs and can only refer to states or particular referential properties (Chafe, 1994). Hence, uses of predicative adjectives often signal a somewhat simplified clause structure and a stative/descriptive type of text.

Constructions with existential *there* also belong among stative features that introduce new content while adding minimal information (Quirk et al., 1985); for example, *There is a huge unemployment in Japan these days* (from a text written by a Japanese NS). Biber (1988) comments that because *there* is employed in syntactically and lexically simple constructions, it is more common in spoken and informal than in written registers. A vast majority of existential *there* constructions occur with *be* as the main verb, predominantly in conversational genres (Biber et al., 1999; Brazil, 1995; Quirk et al., 1985).

Unlike existential *there* structures, however, those with *it*-cleft are considered to be syntactically complex but have relatively little lexical content (Biber, 1988); for example, *It is not necessary for people in this great country of ours to have three SUVs per family* (from a text written by an NS of English). As has been mentioned, they were included in the study primarily because most *it*-clefts occur with *be*-copula verbs. In academic texts, *it* is used to project impartiality, objectivity, and evidentiality when it refers to whole segments of the preceding text (McCarthy, 1994). The increased level of evidentiality leads to more frequent use of *it*-cleft in academic prose and lends implicit authority to the writer's claim by imparting a sense of hedged objectivity (Myers, 1989, 1996). However, in Scollon's (1994) view, the use of *it*-cleft in academic and scientific writing

in English is conventionalized to a large extent and marks the text for a relatively formal register.

Simple Lexical Features

Vague nouns (e.g., *guy, man, people, society, stuff, thing, woman, world*) represent one of the most prominent features of conversational genre and lexically simple prose; for example, *When people grow up, they begin their way in the society, and they have to remember all the things their parents taught them* (from a text written by a Chinese NS). Such nouns are generic (Quirk et al., 1985), and their meanings are rarely homogeneous and almost always depend on the contexts in which they are used. Their most prominent semantic characteristic is that they refer to objects, concepts, and events that are not well defined and have few clear-cut lexical boundaries in the nonlinguistic world (Channell, 1994). Some may actually lack many properties of nouns and function as nominal placeholders in phrases and clauses (e.g., *whoever, whatever*, and other *-ever* words). Based on his corpus analysis, Francis (1994) found that nouns such as *man, thing, stuff*, and *way* are by far the most frequent in popular print media and social interactions. Vague nouns that were included in the frequency counts in their singular and plural forms were, for example, *boy, human(s), human being(s), girl, guy*, nominal *-ever* forms (*whatever, whichever, whoever*), *kid, man, people, person, society, stuff, thing(y), way, whatnot, woman*, and *world* (examples of lexical items here and below are limited to those encountered in both NS and NNS student texts).

Public verbs (Quirk et al., 1985) refer to actions that can be observed publicly and that are used to introduce indirect (and reported) statements (e.g., *argue, explain, promise, say, speak*). An example (from a text written by an NS of English) is *I admit that I like having money, and whoever tells you that they don't is less than honest, I guarantee you*. The verbs in this lexical class are more common in the spoken than in the academic register and are associated with an interpersonal/interactive type of discourse. In general, however, some verbs in this class, such as *ask, say, tell*, and *talk*, are prevalent in conversations and are considered to be particularly lexically simple (Biber et al., 1999; Channell, 1994). Like other lexical features, public verbs were counted in all singular and plural forms. The list of public verbs, which is extensive, includes, for example, *acknowledge, add, admit, agree, announce, ask, assert, bet, claim, comment, complain, confess, confirm, contend, declare, deny, disclose, explain, guarantee, hint, insist, maintain, mention, object, offer, predict, present, protest, remark, repeat, reply, report, say, state, suggest, swear, talk, tell, warn*, and *write*. Many of the less frequent items here and in other lists were encountered in NS and NNS essays in the form of clichés or common colloquial

expressions (e.g., . . . *I guarantee you/it, I confess . . ., he/she swore that . . ., this is true, I swear, disclose private information, protest in vain,* or *I object*).

Private verbs (Quirk et al., 1985) describe or refer to mental states (e.g., *know, learn, think*) and nonobservable intellectual acts that are private, such as emotive acts (*feel, hope*), mental acts (*realize, understand*), and cognitive acts (*believe, conclude, forget, recognize*). An example (from a text written by a Japanese NS) is *When I decided to study film, my father thought that I am completely crazy because there are no jobs in it, but he probably guessed that I knew it.* Private verbs are almost three times more common in the spoken register than public verbs and nearly six times more frequent in conversations than in academic prose (Biber et al., 1999). Overall, in informal conversations, private verbs are practically as common as vague nouns (see above) (Channell, 1994). Some private verbs are *accept, assume, believe, check, conclude, consider, decide, demonstrate, discover, doubt, dream, establish, estimate, expect, feel, find, gather, guess, hear, hold, imagine, imply, indicate, judge, know, learn, mean, note, notice, observe, perceive, presume, pretend, prove, realize, reason, recall, reflect, remember, reveal, see, suppose, suspect, study, think,* and *understand.*

Expecting/tentative/wanting verbs refer to the future time and are often employed in tentative constructions that imply an element of uncertainty (Quirk et al., 1985). An example (from a text written by a Korean NS) is *My brother wanted to go to Australia where his girlfriend went, but my mother tried to introduce him to another girl she liked because she wanted my brother to stay in Korea.* These verbs occur more frequently in speech and informal register than in writing, and it is often difficult to clearly differentiate between these and, for instance, private verbs (Biber, 1988). Expecting/wanting/tentative verbs are relatively rare compared with those in other classes. Johnson (1989) found that in formal academic papers, NNSs used substantially more private and expecting/tentative verbs than NSs did because these verbs are particularly frequent in conversational discourse and social interactions. The most frequent verbs in this lexical class are *attempt, desire, expect, like, plan, try, want,* and *wonder.* The frequency counts included expecting/tentative verbs in all forms.

The Writers

Of the 206 NSs whose essays were analyzed, 89% were first-year students enrolled in required first-year regular composition classes at private universities. The other 11% were similarly enrolled in first-year composition classes in a public university in the Midwest. These students had not received prior writing instruction at the university level, and the

placement and diagnostic tests were administered to the NSs at the very beginning of their first required writing class.

All of the 877 NNSs were international students (NSs of Chinese, Japanese, Korean, Indonesian, and Arabic) who had achieved a relatively high level of L2 proficiency, with TOEFL scores ranging from 543 to 603 (an average of 577). All had been admitted to undergraduate and graduate programs and actively pursued studies toward their degrees in four comprehensive U.S. universities. More than 78% had earned associate's degrees from various community colleges before their admission to 4-year universities, and of the other 22%, 16% were first-year and 6% were graduate students. Successful completion of studies in community colleges and admission as juniors in a 4-year university point to the students' substantial exposure to and experience with U.S. academic frameworks and training. In addition, more than half of the NNS first-year students had been enrolled in U.S. high schools (and boarding schools) for 3–4 years. Because the essay corpus consists simply of placement and diagnostic tests routinely administered to all students, no attempt was made to differentiate the NSs or NNSs by gender or age.

The Prompts and Essays

The prompts for the NS and NNS essays were identical:

1. Some people believe that when parents make their children's lives too easy, they can actually harm their children instead. Explain your views on this issue. Use detailed reasons and examples.

2. Many people believe that grades do not encourage learning. Do you agree or disagree with this opinion? Be sure to explain your answer using specific reasons and examples.

3. Some people learn best when a classroom lesson is presented in a serious, formal manner. Others prefer a lesson that is enjoyable and entertaining. Explain your views on this issue. Use detailed reasons and examples.

4. Many educators believe that parents should help to form their children's opinions. Others feel that children should be allowed to develop their own opinions. Explain your views on this issue. Use detailed reasons and examples.

5. Some people choose their major field of study based on their personal interests and are less concerned about future employment possibilities. Others choose majors in fields with a large number of jobs and options for employment. What position do you support? Use detailed reasons and examples.

The prompts were modeled on those in the Test of Written English, administered by Educational Testing Service, and the Michigan English Language Assessment Battery, as well as those found in many writing/composition textbooks. In such prompts, the intention is to elicit writing samples by providing context based on experiences typical of most young adults entering U.S. universities. All the essay prompts were designed to elicit essays in the rhetorical mode of argument/exposition with the purpose of convincing or informing an unspecified general audience (Hacker, 1994; Leki, 1999). All students were given one class period (50 minutes) to write the essays.

The Corpus

The corpus consists of 322,750 words in 1,083 essays written by speakers of six L1 groups (see Table 1 for the number of essays, the average number of words per essay, and the total number of words in the essays per L1 group). Despite some variability in the numbers of NS and NNS texts, the numbers of essays written for each prompt (see Table 2) are largely comparable and on the same order.

Statistical Analysis

Features were tagged by hand over a period of 2 years by three trained taggers: myself and two assistants. Because the study included only handwritten, in-class essays and because both L1 and L2 student texts often contained numerous misspellings and incorrect morphosyntactic and lexical forms, a consideration was whether to count the features by hand or to type the text and correct the errors to allow for a computer-

TABLE 1
Essays and Words in the Sample by L1 Groups of Writers

L1 group	Essays	Words/essay (*M*)	Words/sample
English	206	285	58,710
Chinese	190	299	56,810
Japanese	184	264	48,576
Korean	166	275	45,650
Indonesian	183	360	65,880
Arabic	154	306	47,124
Total	1,083	292	322,750

TABLE 2
Essays by L1 Group and Prompt

L1 group	Prompt				
	1	2	3	4	5
English	44	36	40	47	39
Chinese	39	39	39	34	39
Japanese	32	35	34	41	42
Korean	32	33	33	32	36
Indonesian	35	35	37	35	41
Arabic	30	32	32	30	30
Total	212	210	215	219	227

ized count of the features. The decision was made to leave essay texts intact and count the features by hand. In tagging features, we did not count errors or omissions but merely counted specific text features as they appeared without second-guessing the writers' intentions. For this reason, no intertagger reliabilities were obtained.

To establish whether NSs and NNSs used the simple syntactic and lexical features in their essay texts similarly, the taggers counted the number of words in each of the 1,083 essays and then counted each feature noted above. For example, NS Essay 2 for Prompt 1 consisted of 250 words and included one vague noun (*people*) and three private verbs (*believe, feel, think*). The percentages of these nouns and verbs in the text were computed (i.e., for the noun, $1/250 = 0.40\%$; for the verbs, $3/250 = 1.20\%$). Percentages were calculated separately for each syntactic and lexical feature in each essay.

Nonparametric statistical comparisons of the NS and NNS data were employed because the majority of the percentage rates were not normally distributed. The Mann-Whitney test was selected as a conservative measure of differences between the NS and NNS data because in some L1 groups only a small number of participants wrote essays in response to a particular prompt. The analysis did not include comparisons of averages because averages often obscure the distribution of frequencies in the sample. For example, if in one NS essay the rate of vague nouns was 7% and in another only 1%, then the mean rate of vague noun uses in these essays would be 4%, which clearly does not reflect the frequency distribution accurately. The ranges are reported to reflect the frequency of use for each feature.

RESULTS AND DISCUSSION

Results indicated significant differences between the structures and lexical forms identified in the native and nonnative texts.

Be-Copula and Attendant Structures

As the data in Table 3 demonstrate, uses of *be*-copula as the main clause verb were significantly more common in NNS than NS texts. In fact, in the essays of Korean speakers the median frequency percentage rates of this feature were almost twice those in the prose of native English speakers, and in texts written by speakers of Japanese and Indonesian, the median rates exceeded those of English speakers by over half the median rate. In many cases, L2 writers relied on *be*-copula and the resulting stative descriptions to advance and support their claims practically to the exclusion of alternative means of essay development.

In Example 1, the NNS writer explains his position on whether parents should allow their children to form their own opinions by means of relatively static exposition:

1. Children's opinions <u>are</u> always a controversial topic. Children <u>are</u> important to their parents, and some believe that older children <u>are</u> mature enough to flourish their opinions without parents' interference. Some argue that children <u>are</u> incapable of distinguishing right from wrong, and therefore it <u>is</u> the responsibility for parents to construct their children's opinions. The two beliefs <u>are</u> contradictory to each other, and many parents <u>are</u> anxious about this dilemma. (native Chinese speaker)

TABLE 3

Median Frequency Rates for *Be*-Copula and Attendant Structures in
NS and NNS Academic Essays (%)

| Marker | L1 group | | | | | |
	English	Chinese	Japanese	Korean	Indonesian	Arabic
be-copula	2.50	3.72**	3.80**	4.12**	3.80**	3.45**
Range	10.71	12.36	11.97	14.81	11.23	8.88
Predicative adjectives	1.68	3.56**	3.54**	3.24**	3.33**	3.13**
Range	8.33	10.83	11.42	20.83	11.98	12.56
there (existential)	0.00	0.00	0.00	0.25	0.00	0.36**
Range	1.70	2.48	3.03	4.07	3.55	2.23
it-cleft	0.54	0.24**	0.36**	0.00**	0.18**	0.29**
Range	3.25	2.04	2.81	1.71	1.37	3.96

$**p \leq 0.05$, two-tailed.

An important observation about this excerpt is that, of the seven occurrences of *be*-copula, five are accompanied by predicative adjectives (*important, mature, incapable, contradictory,* and *anxious*).

In fact, predicative adjectives play an important role in descriptions provided in L2 texts (Table 3): The essays of all five NNS groups contained significantly higher rates of predicative adjectives than those of native English speakers. Specifically, in the prose of all NNS writers, the median frequency rates of predicative adjectives were twice those in NS essays or even higher. The range data further show that in some NNS texts, predicative adjectives represented at least 10% of the total number of words, and in essays of Korean speakers, more 20% of all words. Example 2 illustrates the key role of these simple and stative descriptive features in NNS texts.

> 2. People think that a teacher and students must be <u>serious</u> and <u>diligent</u> in the classroom. I understand this idea, and I think that learning is <u>significant</u> to keep up. In a serious class, people are <u>sleepy</u> or <u>unattentive</u>. If people are <u>unattentive</u> in the class, learning is <u>hard</u> for students. In contrast, if people think that the lesson is <u>enjoyable</u>, most people have interest in the class and the subject. In general, if people have interest in the class and the subject, the lesson that is <u>enjoyable</u> is <u>better</u> than a serious lesson. Also, teachers are usually more <u>friendly</u> in the classroom that is <u>joyful</u> for students than a serious classroom. If teachers are <u>friendly</u>, most students feel more <u>comfortable</u> to ask a question to the teacher. In most cases, if the lesson is <u>serious</u>, the teacher is <u>serious</u>, too. (native Japanese speaker)

The writer constructs her argument that enjoyable classes are better than serious classes by juxtaposing two descriptions that pivot on *be* + predicative adjective constructions. By contrasting the two situations, the text in Example 2 is organized along the lines of *x is good, and y is worse* without actually explaining why this is the case.

On the other hand, NS texts contained significantly lower median frequency rates of *be*-copula as the main-clause verb as well as lower rates of predicative adjectives. In many essays, NSs employed a greater variety of constructions, including activity and causative verbs (Biber et al., 1999). In Example 3, the writer constructs an argument similar to that in Example 2 but employs a different strategy for making her points, focusing on the benefits of serious and entertaining assignments.

> 3. Writing about serious topics can bring out unique and well-planned details that create its setting. Serious classes may form around social occurrences, family affairs, and political events in general. The discussion will educate the individual and allow him or her to mature. . . . While working with serious topics can increase one's maturity, entertainment can bring out the creativity in the person. I find writing about something

I enjoy makes the paper a lot more pleasing to the reader than a serious assignment. Discussing the things I like and enjoy helps me express myself and show who I really <u>am</u>. (native English speaker)

Note that the texts in Examples 2 and 3 are both relatively simple and employ limited ranges of grammar and vocabulary constructions. However, the NS writer of Example 3 relies on a greater number of verbs (*bring out, create, form, educate, increase, express, show*) and fewer stative descriptions. In L2 instruction, it would be hard to argue that the verbs employed in Example 3 are so advanced that they would be difficult to teach (see Conclusions and Implications below).

As mentioned above, syntactic structures with existential *there* often occur with *be*-copula in written academic texts (Biber et al., 1999). The uses of existential *there* are comparatively simple (Biber, 1988; Quirk et al., 1985) and frequent in conversations (Brazil, 1995). The frequency rates of *there* constructions show that at least half the English, Chinese, Japanese, and Indonesian speakers did not employ them in their texts ($M = 0.00$), whereas speakers of Arabic used them significantly more frequently. For example, texts in Examples 4 and 5 contain several uses of existential *there* in discussions of how to choose a major. Both include descriptions of two opposing perspectives without elaboration, similar to Example 2.

4. <u>There</u> are many academic fields and various types of occupations. <u>There</u> are people who think about their personal interests, and other people think about employment opportunities. (native Arabic speaker)

In Example 4, the introductory sentences about many academic fields and two groups of people contain stative generalizations with existential *there* constructions and minimal information outside of that presented in the prompt. In Example 5, the *there* construction that forms a part of the thesis statement also merely lists the points to be covered later in the text without an explanation.

5. When students decide their major fields, they are affected by several factors. Among these factors, <u>there</u> may be interests of students, advice of parents, a good chance to get a job, etc. These effects are different from one person to another. (native Korean speaker)

NS texts, however, exhibited a greater range of syntactic constructions in their expositions. The text in Example 6 is similar to that in Example 5 and includes an introductory excerpt without existential *there* constructions:

6. Choosing a major field of study is a very difficult decision. The factors that influence a person's choice of study can range from financial needs

to simple enjoyment of the area. Although no single factor can determine one's career, personal interests, potential salary, and future job availability, all play a role in this decision. (native English speaker)

It would be difficult to argue that the text in Example 6 is syntactically and lexically sophisticated. However, it displays a greater variety of constructions and verbs (i.e. *range, determine, play a role*) than the texts in Examples 4 and 5. The simplicity of the NS text in Example 6 further emphasizes that L2 writers rely on a particularly limited syntactic and lexical repertoire that may not take a great deal of time and effort to expand.

It-cleft is considered to be an advanced syntactic construction, and not surprisingly it occurred at significantly higher rates in NS texts than in those of any NNS group studied. The median rates of 0.00 for these constructions in the essays of Korean speakers indicate that fewer than half of all writers in these L1 groups employed *it*-cleft in their prose. Furthermore, the frequency rates in the texts of Chinese and Indonesian speakers are less than half the rate in NS essays. The median frequency rates of these structures in NNS texts were not nearly high enough to explain the frequent uses of *be* as the main verb. In many NNS texts, *it*-cleft constructions were largely identified in *it* + copula + adjective patterns, similar to those noted in Shaw and Liu's (1998) sample:

> 7. Everyone knows that people go to school to get an education that will help them to get a job. . . . It is clear that an art major can't pay off their education. (native Chinese speaker)

On the other hand, NS prose contained *it*-cleft constructions with a greater variety of verbs and other attendant elements, such as prepositional phrases and subordinate clauses:

> 8. It is because of their career goals that students actually study. Few people pay thousands of dollars each year just to read textbooks and write papers. (native English speaker)

In Example 8, an *it*-cleft structure is followed by a complex prepositional phrase of reason, introduced by *because of,* and a subsequent subordinate clause, *that students actually study.* Although both NS and NNS essays contained many *it* + copula + adjective constructions, the range of *it*-cleft structures with other types of syntactic elements was greater in NS essays than in NNS prose.

Vague Nouns and Public, Private, and Expecting/Tentative Verbs

The data in Table 4 emphasize that the constraints in NNS writers' lexicon severely limit their options in producing academic essays. The median frequency rates of particularly common and simple vague nouns and public, private, and expecting/tentative verbs in NS texts were significantly lower than those in the essays of all NNS groups studied.

Vague nouns occurred in L2 texts significantly more often than in NS prose for all groups of NNS writers. Example 9 demonstrates the severe limitations in the vocabulary range that to some degree characterize many academic essays produced by NNS students (see also Examples 11 and 13). In fact, the writer relies on just a handful of nouns (*grades, lessons, people, things, class*) to argue his position. As Channel (1994) found in her analysis of a sizable conversational and written corpus, vague nouns, such as *people* and *things*, often function as placeholders in text.

9. <u>People</u> always study hard to get a good grade. <u>People</u> try to get a better grade, as well as they can, and do all the <u>things</u> in class. In my opinion, I disagree with the opinion that said grades to do encourage learning. Because I think that if <u>people</u> get a good grade, it shows that they have learned a lot about this lesson and did all the <u>things</u>. This is an example about grades encourage learning. When <u>people</u> get a bad grade, it means that they haven't learned all about their lessons, so that they fail in the

TABLE 4
Median Frequency Rates for Vague Nouns and Public, Private, and Expecting/Tentative Verbs in NS and NNS Academic Essays (%)

Marker	English	Chinese	Japanese	Korean	Indonesian	Arabic
			L1 group			
Vague nouns	1.48	2.78**	2.20**	3.70**	2.31**	2.61**
Range	8.33	10.83	11.97	14.04	9.23	13.46
Public verbs	0.38	1.58**	1.06*	1.25**	1.04**	1.51**
Range	6.73	7.06	7.24	8.05	8.89	7.89
Private verbs	2.38	3.69**	3.66**	3.24*	3.85**	2.27
Range	10.81	8.60	9.23	10.02	13.33	9.47
Expecting/tentative verbs	0.49	0.87**	1.14**	1.75**	1.02**	0.80**
Range	3.41	6.55	7.64	8.55	5.69	4.76

$*p \leq 0.05$, one-tailed. This test, which uses only one tail of an underlying distribution of values to determine significance, is used for testing a directional (i.e., one-tailed) hypothesis.
$**p \leq 0.05$, two-tailed.

final exam and they must take their class again. Grades are important in learning because from their grades <u>people</u> can see whether they have learned about their lessons or not. (native Indonesian speaker)

In Example 10, in which an NS discusses whether grades encourage learning, the writer's lexicon does not appear to be more advanced than that of a typical first-year composition student. Nonetheless, the lexical range of the text in Example 10 seems to exceed that of the text in Example 9, despite colloquialisms (*millions, to be supposed to, pressures, come up with*) and clichés (*the learning process, as quickly as possible, an average student*).

10. Each year, millions of high school and college students are given grades. These grades are supposed to be a reflection of what the student learned while they were in class. I think grades can encourage learning if they are given properly. Grades allow students to measure how much they have learned and to determine how much more work is needed. Yet, the grading system does provide a significant obstacle because it puts pressures on the teachers and students to come up with the grade. In order to give marks for a student as quickly as possible, the learning process is forced to rush too quickly. Material needs to be covered and learned at a rate that is too fast for an average student. (native English speaker)

Although the text in Example 10 is clearly produced by a writer at a basic level, its range of grammar and lexical structures seems superior to that in Example 9, written by a student whose academic standing exceeded that of the first-year NS, who had had no college-level training.

In addition to vague nouns, the median frequency rates of three lexical classes of verbs indicate the restricted lexical repertoire in NNS academic prose. In particular, the median frequency rates of public verbs were significantly higher in NS essays than in those of Chinese, Korean, Indonesian, and Arabic speakers. The median frequency rates of private verbs were also significantly higher in NS essays than in those of Chinese, Japanese, and Indonesian speakers.

Public, private, and expecting/tentative verbs are very common and lexically simple. Biber et al.'s (1999) extensive analysis of a large English language corpus shows that public and particularly private and expecting/tentative verbs are highly frequent in conversational discourse, and many occur 1–2,000 times per 1 million words.

According to Biber et al. (1999), the semantic content of public verbs usually pertains to communications, and in conversations the most frequent public verbs include *say, tell, talk, ask, write,* and *speak.* In the essays in this study, the median frequency rates of public verb use were significantly higher in the essays of Chinese, Japanese, Korean, Indonesian,

and Arabic speakers than in NS texts. In fact, the frequency rates of public verbs in NNS texts were two to four times as high as those in NS essays.

The text in Example 11 includes a number of public verbs, a few of which are actually unimportant to the writer's argument. However, given that NNS writers often have a restricted range of lexis, it is not particularly surprising that the author resorts to the vocabulary accessible to her.

11. I strongly <u>agree</u> that parents can harm their children. Most of the parents don't understand that giving their children materialistic things, such as money, an expensive watch, and a sports car is not good for children. I can <u>say</u> everything children <u>ask</u> for their parents buy for them. For example, one of my friends always <u>asks</u> her parents to buy her the brand name fashion and daily necessities which are so expensive. . . . Parents don't <u>talk</u> to their children and <u>explain</u> to them how they should do something because parents already made the arrangements to take care of everything. . . . Teachers sometimes <u>talk</u> to spoiling parents, but no good comes out of it. (native Chinese speaker)

The NS text in Example 12 similarly argues that parents may harm their children by spoiling them. Although the NS text seems to be basic and unsophisticated, its syntactic and lexical range is greater than that in Example 11. In the 120-word excerpt shown in Example 12, the writer employs only one public verb (*ask*).

12. A friend of mine has everything she wants, and maybe, this is because she is an only child. If the parents provide for the child's every need and whim and if the child never has to achieve something on their own, they may not learn the value of hard work. Or this young person may find the "real world" is not what they are used to. It may also be harmful to the child as an adult to find themselves unprepared for reality. Later in life, as an adult, this former child may regret all the help his/her parents provided and <u>ask</u>, if my parents loved me, why didn't they help me to be more prepared for success. (native English speaker)

Although NS students included many public and private verbs in their texts, the median frequency rates of private verbs (*believe, feel, learn, study, think, understand*) in NS essays were also substantially lower than in the texts produced by Chinese, Japanese, Korean, and Arabic speakers. As mentioned, private verbs are associated with mental and emotive states that are typically found in relatively personal and static descriptions (Quirk et al., 1985). According to Biber et al. (1999), in conversational discourse the most frequent private verbs include *see, know, think, find, feel,* and *remember*.

The number of lexical, ideational, and syntactic redundancies and the syntactic and lexical simplicity of the L2 text in Example 13 may exasperate the reader. The verb *learn* is repeated five times; *understand* and *think,* twice each; and *forget, feel, see,* and *remember* each occur once.

13. People can <u>learn</u> in many different ways. Some people <u>learn</u> well when the subject is presented in a situation that is serious. I <u>learn</u> well in a class that is formal because I can <u>think</u> more deeply about the subject, and it helps me to <u>remember</u> the material easier. Then I won't <u>forget</u> it easily and <u>think</u> about it for a long time. My sister told me that when she <u>learns</u> something, and there are visual aids beside the information, she can understand the subject better. I <u>feel</u> that I can <u>understand</u> the subject better in one way, and my sister can <u>see</u> it better in another. This is because different people <u>learn</u> in different ways. (native Indonesian speaker)

Although some researchers have commented that NNS students often do not include sufficient elaborations and detailed descriptions in their texts (Lee & Scarcella, 1992), the reason may lie partly in the fact that many L2 writers do not have the vocabulary range to produce elaborations. For instance, the lexis employed in Example 13 may be proximate to that of NS school-age learners, who simply write down their experiences within the confines of their linguistic repertoire (Bereiter & Scardamalia, 1987).

On the other hand, the first-year NS writer in Example 14 also advocates a serious approach to classroom teaching but without relying on private verbs. Although the text in also contains redundancies (*students, time, lecture, given*), they do not seem nearly as distracting as those in Example 13 because of the presence of variations in lexis, such as *lecture/lesson/demonstration/presentation, effective/appropriate,* and *interactive/engaging/entertaining.* This variation in lexis with similar semantic content is prominently absent from L2 texts.

14. Each individual has their personal learning style. . . . For me, a serious lecture or lesson may be most effective with higher level students. If a student has very little time, they may <u>feel</u> the need to get the maximum amount of information in an hour and <u>feel</u> that demonstrations or games, meant to be interactive and engaging are wasting their time. A serious lecture may also be more appropriate when the information given represents important facts or analysis. Highly entertaining presentations have a high chance of taking away from the message that is being given. (native English speaker)

Another feature that noticeably separates NS and NNS texts is the presence of common idiomatic expressions such as those found in

Example 14 (e.g., *have time, waste time, a high chance, take away from*). Even though these expressions may seem simple, they are not very common in NNS academic texts.

Like public and private verbs, expecting/tentative verbs occur quite frequently in conversations. Biber et al.'s (1999) study shows that, in the conversational corpus, the verbs *want, like,* and *try* occur 500–2,000 times per 1 million words. In the university essays examined in this study, the median frequency rates of expecting/tentative verbs in NNS prose were two to three times those in NS texts. For example, the median frequency rate in the NS essays (0.49) was approximately half of that in the essays of Indonesian speakers (1.02) and less than a third of that in the texts of Korean speakers (1.75). On the whole, in this study, the median frequency rates for expecting/tentative verbs in NS essays were significantly higher than those in the essays of NNSs in all L1 groups.

In Example 15, the NNS writer explains the differences between appropriate learning activities in the United States and her country, using six expecting/tentative verbs—*try, want* (used twice), *like* (used twice), and *plan*—in a 77-word excerpt.

15. My classmates <u>try</u> to get the most of their classes. They <u>want</u> to know how something happens because they <u>like</u> to talk to other people during class. In my country, students <u>want</u> to get university degrees, but they do not talk in groups. When I <u>planned</u> to go to the U.S. to study, I didn't know that I would have to talk in class in front of other people, and I don't <u>like</u> this way of teaching. (native Chinese speaker)

Note that the NNS writer explains his position based largely on personal preferences and expectations of particular classroom activities conveyed by means of expecting/tentative verbs, such as *like* and *want.* As mentioned above, in many cases NSs commented on both the pros and the cons of the issue at hand in their essays. In Example 16, the NS presents his argument and accounts for possible objections to his position.

16. My personal preference is to have interactive classes because I can learn about my presentation style and improve it. On the other hand, many students avoid speaking up in class because they do not have enough confidence. It basically depends on the learning preferences of the student. I differ from those students in the way that I prefer learning together with others in a group or during a discussion. (NS)

The text in Example 16 covers a broader linguistic range, including such nouns as *preference, style,* and *confidence,* which, taken together with other lexis, create a more sophisticated impression than the counterposition in Example 15, which is expressed by means of expecting/tentative verbs, such as *like* and *want.*

On the whole, the combined use of various constructions, such as *be*-copula and simple nouns and verbs, all of which are highly frequent in conversational rather than written discourse, lead to the appearance of general syntactic and lexical paucity in NNS prose. None of these features alone can project a sense of simplicity, but taken together, and with a preponderance of simple, high-frequency lexical items, such as *people, thing, say, think, and like/want,* these characteristics can be quite damaging when it comes to assessment and evaluation of university-level essays (Ferris & Hedgcock, 1998; Vaughan, 1991).

CONCLUSIONS AND IMPLICATIONS

A detailed analysis of L2 academic essay texts provides clear evidence that NNS students with a relatively high academic standing employ significantly higher median rates of simple syntactic and lexical features than newly admitted first-year NS students do. An implication of this finding is that the NNSs' productive range of grammar and lexis is comparatively small and consists largely of constructions prevalent in spoken and conversational discourse as well as high-frequency, everyday vocabulary items. Although these findings may not be particularly surprising for basic learners or new arrivals, they are disheartening in the case of NNS students who have spent years pursuing academic studies in the United States.

Specifically, the study identifies as particularly prevalent the use of *be*-copula as the main verb most often associated with employment of predicative adjectives, as well as frequent use of vague nouns and public, private, and expecting/tentative verbs. On the other hand, the frequency rates of academic and formal *it*-cleft constructions (Scollon, 1994) are markedly low in L2 texts. Although there is little doubt that the NNS students in this study had been exposed to L2 academic reading and text for comparatively lengthy periods, mere exposure to academic text and reading may not be sufficient for L2 learners to attain the advanced academic proficiency essential for success in their academic endeavors. Thus, instruction for university-bound L2 students needs to concentrate on expanding their syntactic and lexical repertoire.

A recent study by Norris and Ortega (2000) undertakes probably the most comprehensive analysis of published data on the value of grammar instruction. The authors state that in many cases, it is not easy to tell whether communicative, explicit, or meaning-focused instruction led to greater degrees of L2 learning and acquisition because of the disparate sample sizes and statistical analyses employed in various research studies and publications. Thus, to make sense of research findings published in the past two decades, Norris and Ortega standardized the results of 49

studies on L2 learning, acquisition, and grammar instruction. The outcomes of their substantial undertaking show clearly that in L2 teaching, "focused instructional treatments of whatever sort far surpass non- or minimally focused exposure to the L2" (p. 463). Meeks (1994) emphasizes that academic L2 writers need to be taught the uses of advanced textual features, such as grammar and lexis, because without such instruction students may not learn to recognize distinctions between formal written and spoken registers. She underscores that, particularly in academic and professional settings, students need to be taught the "tools" (p. 41) that can help them increase their syntactic and lexical repertoire.

One potentially effective teaching technique for working with the overused *be*-copula and with public, private, and expecting/tentative verbs, all of which are frequently encountered in conversational (and informal) discourse, is to give learners lexical alternatives, similar to those found in texts of basic NS writers. For example, to circumvent repeated uses of the noun *people,* a writer might substitute the lexically and contextually proximate vocabulary items *individuals/adults* (or *children/youngsters*) and, in academic texts based on students' educational experience, *students/learners/audience.* Furthermore, alternatives to the excessively popular verb *think* might be *consider/debate/deliberate/explore,* and *like/want* can be easily replaced by *prefer/value/appreciate.* Few of these lexical options are particularly sophisticated, nor do they need to be to provide textual and lexical variety. Activities focused on uses of lexical alternatives may not be hard to develop and implement.

Instruction on the syntactic and textual properties of predicative and attributive adjectives can deal with adjective shifts in sentences (e.g., *students are careful—careful students*). Teachers can demonstrate how moving adjectives from the predicative to the attributive position can free up half of a sentence and allow writers to express their ideas compactly.

The preponderance of stative (and static) *be*-copula as the main verb in NNS texts further points to a conclusion that NNS students' lexicon may notably lack what Biber et al. (1999) and Brazil (1995) call *existence verbs.* In academic texts, important (but not necessarily complex) verbs in this class not encountered in NNS student texts include *concern, constitute, deal (with), define, derive, deserve (attention, effort), fit, illustrate, imply, lack, matter, possess, relate, remain, reveal, suit, sound, tend,* and *vary.* Other underused verbs in the activity class that may merit attention are *accompany, advance, combine, control, encounter, engage, exercise, expand,* and *explore.* Lists such as these may not be particularly difficult to obtain (e.g., "A University Word List," in Nation, 1990, includes all words common in introductory-level textbooks across several disciplines; see also the other corpus-based studies cited in this article). Working with such items in

somewhat contextually cohesive groupings (e.g., *study, improve, practice, work on,* or *bring about, cause, change, remain, increase, decrease*) to construct short pieces of academic discourse, such as reports or literature reviews, is not very complicated.

In light of the many studies of syntactic and lexical features in academic and other types of texts carried out in the past decade, the fact that essays of experienced NNS students lack many characteristics of academic text largely speaks to the shortfalls of L2 teaching and learning. The reasons that L2 instruction often does not meet its stated goals of preparing learners for their academic objectives may be complex, but some may have to do with choices in L2 teaching methodologies and popular methods widely adopted in teacher education. On the whole, teachers of academically bound students and researchers of academic prose may find it fruitful and constructive to find out how to improve students' text production skills to yield more sophisticated syntactic constructions and lexis so that the students are at a smaller disadvantage when they leave the ESL classroom.

THE AUTHOR

Eli Hinkel has taught ESL and applied linguistics as well as trained teachers for more than 20 years and has published books and numerous articles on learning second culture and L2 grammar, writing, and pragmatics. She is also the series editor of Lawrence Erlbaum's ESL & Applied Linguistics Professional Series.

REFERENCES

Adams, V. (2001). *Complex words in English.* Harlow, England: Pearson.

Allen, V. F. (1983). *Techniques in teaching vocabulary.* Oxford: Oxford University Press.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition.* Hillsdale, NJ: Erlbaum.

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

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

Brazil, D. (1995). *A grammar of speech.* Oxford: Oxford University Press.

Carlson, S. (1988). Cultural differences in writing and reasoning skills. In A. Purves (Ed.), *Writing across languages and cultures: Issues in contrastive rhetoric* (pp. 109–137). Newbury Park, CA: Sage.

Chafe, W. (1994). *Discourse, consciousness, and time.* Chicago: University of Chicago Press.

Channell, J. (1994). *Vague language.* Oxford: Oxford University Press.

Cumming, A., Kantor, R., Powers, D., Santos, T., & Taylor, C. (2000). *TOEFL 2000 writing framework: A working paper* (Monograph Series No. 18). Princeton, NJ: Educational Testing Service.

Davidson, F. (1991). Statistical support for training in ESL composition rating. In L. Hamp-Lyons (Ed.), *Assessing second language writing* (pp. 155–165). Norwood, NJ: Ablex.

Ferris, D., & Hedgcock, J. (1998). *Teaching ESL composition*. Mahwah, NJ: Erlbaum.

Francis, G. (1994). Labelling discourse: An aspect of nominal-group cohesion. In M. Coulthard (Ed.), *Advances in written text analysis* (pp. 83–101). New York: Routledge.

Frase, L., Faletti, J., Ginther, A., & Grant, L. (1999). *Computer analysis of the TOEFL Test of Written English* (Research Report No. 64). Princeton, NJ: Educational Testing Service.

Hacker, D. (1994). *The Bedford handbook for writers* (4th ed.). Boston: Bedford.

Hamp-Lyons, L. (1991). Scoring procedures for ESL contexts. In L. Hamp-Lyons (Ed.), *Assessing second language writing* (pp. 241–277). Norwood, NJ: Ablex.

Hunston, S., & Francis, G. (1996). *Pattern grammar: A corpus driven approach to the lexical grammar of English*. Amsterdam: John Benjamins.

Johnson, D. (1989). Politeness strategies in L2 written discourse. *Journal of Intensive English Studies, 3*, 71–91.

Johnson, D., & Roen, D. (1989). Introduction. In D. Johnson & D. Roen (Eds.), *Richness in writing* (pp. 1–15). New York: Longman.

Jordan, R. (1997). *English for academic purposes*. Cambridge: Cambridge University Press.

Lee, C., & Scarcella, R. (1992). Building upon Korean writing practices. In F. Dubin & N. Kuhlman (Eds.), *Cross-cultural literacy* (pp. 143–161). New York: Pearson.

Leech, G., Rayson, P., & Wilson, A. (2001). *Word frequencies in written and spoken English*. Harlow, England: Pearson.

Leki, I. (1991). *Understanding ESL writers*. Portsmouth, NH: Boynton/Cook.

Leki, I. (1999). *Academic writing: Techniques and tasks* (3rd ed.). New York: Cambridge University Press.

McCarthy, M. (1994). It, this, and that. In M. Coulthard (Ed.), *Advances in written text analysis* (pp. 266–275). New York: Routledge.

Meeks, K. (1994). Is grammar a dirty word? *College ESL, 4*(2), 39–42.

Myers, G. (1989). The pragmatics of politeness in scientific articles. *Applied Linguistics, 10*, 1–35.

Myers, G. (1996). Strategic vagueness in academic writing. In E. Ventola & A. Mauranen (Eds.), *Academic writing* (pp. 1–18). Amsterdam: John Benjamins.

Nation, I. S. P. (1990). *Teaching and learning vocabulary*. New York: Newbury House.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Norris, J., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*, 417–528.

Ogden, C. (1930). *Basic English: A general introduction with rules and grammar*. London: Paul Treber.

Palmer, H., West, M., & Faucett, L. (1936). *Interim report on vocabulary selection for the teaching of English as a foreign language* (Report of the Carnegie Conference). London: King.

Pica, T. (1985). Linguistic simplicity and learnability: Implications for language syllabus design. In K. Hyltenstam & M. Pienemann (Eds.), *Modelling and assessing second language acquisition* (pp. 137–153). San Diego, CA: College Hill Press.

Pienemann, M. (1985). Learnability and syllabus construction. In K. Hyltenstam & M. Pienemann (Eds.), *Modelling and assessing second language acquisition* (pp. 23–77). San Diego, CA: College Hill Press.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. New York: Longman.

Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.

Reid, J. (1993). *Teaching ESL writing*. Englewood Cliffs, NJ: Prentice Hall.

Reynolds, D. (1995). Repetition in nonnative-speaker writing: More than quantity. *Studies in Second Language Acquisition, 17,* 185–209.

Reynolds, D. (2001). Language in the balance: Lexical repetition as a function of topic, cultural background, and writing development. *Language Learning, 51,* 437–476.

Richards, 1. (1943). *Basic English and its uses.* London: Kegan Paul.

Scollon, R. (1994). As a matter of fact: The changing ideology of authorship and responsibility in discourse. *World Englishes, 13,* 33–46.

Shaw, P., & Liu, E. T.-K. (1998). What develops in the development of second language writing. *Applied Linguistics, 19,* 225–254.

Silva, T. (1993). Toward an understanding of the distinct nature of L2 writing: The ESL research and its implications. *TESOL Quarterly, 27,* 657–676.

Sinclair, J. (1991): *Corpus, concordance, collocation.* Oxford: Oxford University Press.

Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words.* New York: Columbia University, Teachers College.

Vaughan, C. (1991). Holistic assessment: What goes on in the raters' minds? In L. Hamp-Lyons (Ed.), *Assessing second language writing* (pp. 111–126). Norwood, NJ: Ablex.

West, M. (1953). *A general service list of English words.* London: Longman.

Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity.* Honolulu: University of Hawaii Press.