Should We Use Characteristics of Conversation to Measure Grammatical Complexity in L2 Writing Development?

DOUGLAS BIBER AND BETHANY GRAY
Northern Arizona University
Flagstaff, Arizona, United States

KORNWIPA POONPON
Khon Kaen University
Khon Kaen, Thailand

Studies of L2 writing development usually measure T-units and clausal subordination to assess grammatical complexity, assuming that increased subordination is typical of advanced writing. In this article we challenge this practice by showing that these measures are much more characteristic of conversation than academic writing. The article begins with a critical evaluation of T-units and clausal subordination as measures of writing development, arguing that they have not proven to be effective discriminators of language proficiency differences. These shortcomings lead to the question of whether these measures actually capture the complexities of professional academic writing, and if not, what alternative measures are better suited? Corpus-based analyses are undertaken to answer these questions, investigating 28 grammatical features in research articles contrasted with conversation. The results are surprising, showing that most clausal subordination measures are actually more common in conversation than academic writing. In contrast, fundamentally different kinds of grammatical complexity are common in academic writing: complex noun phrase constituents (rather than clause constituents) and complex phrases (rather than clauses). Based on these findings, we hypothesize a sequence of developmental stages for student writing, proposing a radically new approach for the study of complexity in student writing development.

doi: 10.5054/tq.2011.244483

As a reader, your initial reaction to the question posed in the title of this article might have been “No, of course not. What a ridiculous suggestion!” We agree with that reaction. But surprisingly, current practice in the evaluation of L2 writing development focuses primarily on grammatical features that are more prevalent in conversation than in
professional academic writing. Our primary goal in this article is to challenge this practice: we first survey current approaches to the study of complexity in writing development, showing how they rely on conversational grammatical characteristics, and then we propose an alternative set of grammatical features that are more suitable for this purpose.

At least since the 1930’s, researchers in writing development have focused on grammatical complexity, studying how students’ language increases in complexity as those students become more proficient writers. Early research of this type considered mostly the writing of primary and secondary school students who were native speakers of English (e.g., Anderson, 1937; Frogner, 1933; LaBrant, 1933). This focus continued through the 1960’s (e.g., Hunt, 1965; Loban, 1963) but was then extended to the writing of college students in the late 1960’s and 1970’s (see, e.g., Faigley, 1980; Hiatt, 1978; Hunt, 1970; Jakobovits, 1969; Lunsford, 1978).

During that same period, composition teachers and researchers became interested in the writing-as-process approach to writing instruction (see the survey of research in Grabe & Kaplan, 1996, pp. 84–112). This paradigm shift dramatically influenced the focus of first-language writing development research, so that by the 1990’s there were almost no new studies that analyzed the grammatical characteristics of written texts produced by students in traditional (L1 English) composition courses. Rather, writing development research shifted to consideration of the writing process or, more recently, to a focus on student identities, critical thinking, and the larger sociocultural context of writing.

However, at the same time that composition researchers in rhetoric were moving away from the linguistic study of student texts, other researchers were becoming interested in L2 writing development with an overt focus on the linguistic structures used in student texts (see, e.g., Cooper, 1976; Ferris & Politzer, 1981; Flahive & Snow, 1980; Gipps & Ewen, 1974). This trend has continued to the present time, so that it is common now to find second language researchers who focus on

1The term complexity has been employed in many distinct ways within different subfields of linguistics. For example, within psycholinguistics, processing complexity is often measured by the amount of time required to understand a linguistic structure. Within typological linguistics, complexity is often tied to the phonemic and morphological inventory of a language, so that languages (or varieties) with more phonological/morphological distinctions are more complex than languages that make fewer distinctions (see, e.g., the debates presented in the commentary articles published in the journal Linguistic Typology (2001; Issue 2/3 of Volume 5). Within applied linguistics, there is a long and extensive history of using the term complexity to refer to the more advanced grammatical structures that students exhibit as they progress in their language proficiencies. Grammatical complexity has been especially studied in relation to writing development, and it is this use of the term that provides the focus for the present study.
measures of fluency, accuracy and complexity in second language writing (as in the title of the 1998 book by Wolfe-Quintero, Inagaki, & Kim; see, e.g., Brown, Iwashita, & McNamara, 2005; Ellis & Yuan, 2004; Larsen-Freeman, 2006; Nelson & Van Meter, 2007).

MEASURING GRAMMATICAL COMPLEXITY IN PREVIOUS WRITING DEVELOPMENT RESEARCH

Across these decades and subdisciplines, when writing development research has focused on the linguistic description of student texts, one of the key concerns has been the analysis of grammatical complexity. These studies have adopted a production perspective on complexity, based on the expectation that written language production increases in grammatical complexity as language skills develop and students (both native speakers and second language (L2) learners) become more proficient writers. (In contrast, a processing or comprehension perspective on complexity would be more relevant for reading research.)

Most previous research on writing development has adopted a deductive approach, beginning with an a priori definition of grammatical complexity as elaborated structures added on to simple phrases and clauses (see, e.g., Purpura, 2004, p. 91; Willis, 2003, p. 192). Specifically, most studies of L2 writing development have relied on quantitative variables that measure the average length of structural units or the extent of clausal subordination, assuming that longer units and more subordination reflect greater complexity. Student writing development is then assessed by these measures.

The vast majority of these studies have relied on the construct of the T-unit: a main clause and all associated dependent clauses. Two specific measures have been especially popular: mean length of T-unit (MLTU), which relies on the overall length in words of the T-unit, averaged across all T-units in a text, and clauses per T-unit (C/TU), which relies on the number of dependent clauses per T-unit, again averaged across all T-units in a text. For example, the following sentence comprises a single T-unit that is relatively short (11 words) but includes two embedded dependent clauses:

*I don’t know [why I was expecting [to see something else]].*

The reliance on clausal subordination and T-unit–based measures is documented by Wolfe-Quintero, Inagaki, and Kim (1998), who provide an extensive survey of research on L2 writing development through the late 1990’s (see especially Chapter 4). In fact, in their concluding chapter, Wolfe-Quintero et al. single out clauses per T-unit and
dependent clauses per independent clause as the “best [...] complexity measures so far” (pp. 118–119). Perhaps in part because of this recommendation, studies of L2 writing development since 1998 have continued to rely heavily on measures based on the T-unit (e.g., Brown, Iwashita, & McNamara, 2005; Ellis & Yuan, 2004; Larsen-Freeman, 2006; Nelson & Van Meter, 2007), or related measures based on the frequency of subordinate clauses (e.g., Brown et al., 2005; Li, 2000; Norrby & Håkansson, 2007). Ortega (2003) provides strong confirmation that current research continues to employ these same two measures, based on a meta-analysis of empirical research on grammatical complexity in college-level ESL/EFL writing. Of the 27 studies included in her survey, 25 rely on the MLTU to measure grammatical complexity, while 11 used the related measure of C/TU. No other measure was used widely across these studies.2

Surprisingly, despite their widespread acceptance, there is little empirical evidence that T-unit measures and dependent clause measures are appropriate for the assessment of writing development. This shortcoming was noted in a relatively early study by Bardovi-Harlig (1992), who writes “in evaluating the syntactic complexity of compositions written by advanced adult second language learners, T-unit analysis does not seem to reflect accurately the knowledge of the learner” (p. 391). More recently, scholars like Rimmer (2006, 2008), Ravid (2005), Ravid & Berman (2010), and Norris and Ortega (2009) have also problematized the application of subordination-based measures in the study of writing development.

The continuing reliance on these measures to assess writing development reflects traditional notions of complexity and the widespread belief that academic written discourse is complex in that it relies heavily on elaborated structures. For example,

Students [writing chemistry lab reports] engage in elaborated discourse with a high degree of specificity [...] Once they have focused on salient data and evidence, elaborated forms of discourse arrange information into more complex and explicit representations reflective of canonical scientific ideas. (Wright, 2008, p. 292)

[...] in academic writing [...] elaborated structures are generally preferred as they facilitate the readers’ understanding of the text. (Hyland & Tse, 2005, p. 127)

2 Although there have been far fewer empirical studies of language development in speech, recent studies have adopted these same measures to investigate grammatical complexity. Thus, studies like Mehnert (1998), Skehan and Foster (1999), Robinson (2001), and Yuan and Ellis (2003) all rely on measures of subordination, usually average T-unit length or the number of subordinate clauses per T-unit.
In sum, it has been standard practice to assume that students follow a natural progression from simple clause structures to the more complex and elaborated clause structures that are supposedly typical of professional academic writing.

**PREVIOUS LINGUISTIC RESEARCH ON GRAMMATICAL COMPLEXITY IN SPEECH VERSUS WRITING**

Linguists who have studied the grammatical characteristics of spoken and written discourse provide a completely different perspective on complexity, arguing that a dense use of clausal subordination is *not* typical of advanced academic writing. In fact, clausal subordination is much more prevalent in conversational discourse than in academic writing. In contrast, linguistic analyses of written academic texts show that they are composed primarily of embedded noun phrases and prepositional phrases, with comparatively few embedded dependent clauses.

Thus, as early as 1960, Rulon Wells argued that nouns are more important than verbs in academic writing, describing the nominal style of written discourse contrasted with the verbal style of speech. Multidimensional studies of register variation, first undertaken in the 1980s (Biber, 1985, 1986), have used large-scale corpus analyses to document how clausal subordination is typical of speech, while academic writing relies on phrasal modifiers instead of dependent clauses. For example, Dimension 1 in the original multidimensional study of English (Biber, 1988, pp. 104–108) showed that finite dependent clauses—including *that* clauses, WH clauses, causative adverbial clauses, and conditional adverbial clauses—are characteristic of interpersonal spoken registers. In contrast, noun-modifying phrasal features (e.g., attributive adjectives and prepositional phrases) are especially characteristic of formal written registers. A multidimensional study of discourse complexity (Biber, 1992) confirms these differences, while Biber, Conrad, Reppen, Byrd, and Helt (2002; see also Biber, 2006) show that similar discourse patterns distinguish spoken university-level classroom teaching from written university textbooks. Biber, Johansson, Leech, Conrad, and Finegan (1999) provide more detailed descriptions of the grammatical features that are common in conversation versus those that are common in academic writing, showing that most finite

---

3Biber (1995) shows that these patterns seem to hold cross-linguistically, based on a comparison of multidimensional analyses for English, Somali, Korean, and Tuvaluan. Thus, across these four languages, “relative clauses, and nominal modifiers generally, are characteristic of literate registers... In contrast, adverbial subordination is used most commonly in oral registers... Complement clauses and infinitives occur frequently in both oral and literate registers...” (p. 263).

MEASURING GRAMMATICAL COMPLEXITY 9
dependent clause types are considerably more common in speech than in writing.

Halliday (1989, 2004), taking a more theoretical perspective, has also argued that the complexities of speech are dramatically different from those of academic writing and, specifically, that the major grammatical complexities of speech involve dependent clauses, while writing relies on nouns and nominalizations (see, e.g., Fang, Schleppegrell, & Cox, 2006; Halliday & Martin, 1993/1996; Halliday & Mathiessen, 1999). For example,

Spoken language is more intricate than written. [...] From that point of view, it will appear that spoken language is more complex than writing (Halliday, 1989, p. 62)

The highly information-packed, lexically dense passages of writing often tend to be extremely simple in their grammatical structure (Halliday, 1989, p. 87) [...] the complexity of written language is lexical, while that of spoken language is grammatical (Halliday, 1989, p. 63)

[…] something that would in spoken English be typically expressed as a clause is expressed instead [in writing] as a group of words centring on a noun (Halliday, 2004, p. 171).

Thus, from both empirical and theoretical perspectives, it is well established that speech and writing are grammatically complex in dramatically different ways. More important, it is well established that the grammatical features stereotypically associated with complexity—clausal subordination features—are actually much more common in conversation than in academic writing. Thus, if we focused on only clausal subordination features, we would be forced to conclude that conversational discourse is more complex than academic writing. In contrast, the complexities of academic writing are phrasal rather than clausal (see below for a much fuller discussion of these differences).

However, applied linguists seem generally unaware of this body of research. In particular, two stereotypes persist:

1. that grammatical complexity is best measured by consideration of clausal subordination; and
2. that academic writing is obviously more complex than conversation with respect to those features.

The large body of research on L2 writing development (surveyed above) shows that these stereotypes are deeply entrenched: most researchers unquestioningly apply clausal subordination measures to evaluate writing development, never considering the possibility that those measures are actually more characteristic of speech than writing. But the influence of these stereotypes is wider, making it difficult for many
applied linguists across subdisciplines to even imagine alternative possibilities.

To be completely explicit, we are directly challenging both of the stereotypes listed above. With respect to the second stereotype, we show in the following sections that conversation is more complex than academic writing, if we consider only clausal subordination measures (following the practice of most L2 writing research). However, with respect to the first stereotype, we argue that alternative grammatical characteristics (associated with complex noun phrases rather than embedded clauses) are much more appropriate measures of grammatical complexity in academic writing.

OVERVIEW OF THE PRESENT STUDY

Our first goal in the present article is to undertake a critical evaluation of T-unit and subordination-based measures of complexity, arguing that they have not proven to be effective discriminators of language proficiency differences, and that they are not well motivated from a linguistic perspective. This leads us to question the underlying assumption: that extensive subordination is an important measure of grammatical complexity in academic written discourse. In contrast to previous writing development research, which assumed an a priori definition of complexity as structural elaboration, the present article undertakes empirical research to inductively identify the grammatical features that are most strongly characteristic of advanced academic writing.

For these purposes, we undertake a large-scale corpus-based analysis, investigating the distribution of 28 grammatical features in academic research articles, contrasted with the patterns of use in conversation. The analysis is based on a wide range of grammatical devices associated with complexity and used to add elaborating information in written texts, including both dependent clauses as well as phrasal modifiers. The results are surprising: the clausal subordination features that have been traditionally used to assess complexity in writing development are not actually characteristic of professional written discourse. In fact, many of these features are more common in conversation than in academic writing. In contrast, the analysis shows that fundamentally different kinds of grammatical complexity are common in academic writing. The findings have immediate implications for the study of writing development, and we explore those in the conclusion, hypothesizing a series of developmental stages for the use of complex grammatical features associated with advanced writing.
The association between clausal subordination and grammatical complexity is deeply entrenched in linguistic theory. For example, by definition, a simple clause has only a subject, verb, and object or complement. A simple noun phrase has a determiner and head noun. Additions or modifications to these patterns result in complex grammar, with the implicit understanding that more additions result in more complexity. In particular, linguists from several theoretical backgrounds have singled out dependent clauses as one of the most important types of grammatical complexity (often described in contrast to simple clauses or coordinated clauses; see, e.g., Carter & McCarthy, 2006, p. 489; Huddleston, 1984, p. 378; Purpura, 2004, p. 91; Willis, 2003, p. 192).

Reflecting these same general assumptions, Hunt (1965) proposed the T-unit as an omnibus measure of grammatical complexity for the study of student writing development. The logic here is straightforward: as we add more structures to a simple clause, that clause becomes increasingly complex, and the T-unit length becomes longer. Variants of this measure focus on dependent clauses: adding dependent clauses to a simple clause results in a more complex structure, reflected by a greater number of clauses per T-unit.

As noted in the last section, most researchers who study or assess L2-English writing development have relied on the T-unit and dependent clause measures, and those measures have been strongly endorsed by researchers like Wolfe-Quintero et al. (1998, p. 118). Surprisingly, though, there is little empirical evidence to recommend the use of T-unit–based measures for the study of grammatical complexity. Criticisms of the T-unit can be grouped under two general domains:

1. its lack of utility in testing applications
2. its poor theoretical linguistic basis

With respect to the first criticism, it has certainly not been demonstrated that T-unit–based measures are useful for distinguishing among learner groups at different proficiency levels. In fact, just the opposite is the case.

For example, Wolfe-Quintero et al. (1998, pp. 82–86) survey studies of grammatical complexity across developmental levels, most of which used T-unit–based measures. While some studies show improvement across developmental levels, many other studies actually show a decline.

---

4 Hunt (1965) also explored the use of numerous more specific grammatical characteristics as indicators of complexity, but subsequent researchers have usually disregarded his other research and focused almost exclusively on the single measure of the T-unit.
Summarizing the overall pattern for dependent clauses per T-unit (the “T-unit complexity ratio”), Wolfe-Quintero et al. note that “seven studies found a significant relationship between proficiency and the T-unit complexity ratio, while eleven did not” (p. 85). That is, 61% of these studies failed to find a significant relationship between proficiency and complexity as measured by the use of subordinate clauses in T-units.

Ortega (2003) similarly fails to find support for the utility of T-unit–based variables as measures of language proficiency. For example, Ortega analyzes 68 specific comparisons across proficiency levels (from the 27 studies in her sample); some of these are comparisons between adjacent groups, whereas others are comparisons between the lowest and highest proficiency groups (see p. 504). Figure 2 in her study (p. 505) plots the observed difference between proficiency groups for MLTU. Forty-three of the 68 comparisons (≈65%) showed almost no difference for the MLTU across proficiency levels (a difference smaller than ±1.8 words). Only three of those comparisons were reported to be statistically significant.

Thus, T-unit–based measures have not been reliable indicators of proficiency-level differences. In fact, more often than not, empirical studies have failed to find consistent increases for T-unit–based measures as students advance in language proficiency.5 These studies begin with the assumption that the measurement of complexity is not in itself controversial: that we somehow already know that more dependent clauses represents more complexity. However, the uncritical acceptance of this assumption leads to the mysterious conclusion that student writing fails to increase (and often decreases) in complexity as students advance in proficiency. It is not at all clear how to interpret that repeated finding (see, e.g., the discussion in Ortega, 2003, pp. 512 ff., which focuses on differences in instructional settings and proficiency sampling criteria).

In the present article, we challenge the underlying linguistic assumptions of this line of research: we argue that student writing probably does become more complex at higher proficiency levels, but T-unit–based measures and simple subordination measures are not able to capture those complexities. That is, T-unit–based measures confound fundamentally different kinds of grammatical structures, and as a result, it is not surprising that developmental studies have failed to observe consistent increases with respect to these measures.

T-unit–based measures assume a single cline of phenomena: simple versus complex. The simple pole of this cline is uncontroversial: clauses with only a single verb phrase, no dependent clauses, and no clausal or

---

5 A separate issue concerns the use of T-unit–based measures to determine proficiency level in testing applications; the empirical research surveyed in Wolfe-Quintero et al. (1998) and Ortega (2003) suggests that this application is highly suspect at best.
nominal modifiers. The problem comes in defining the complex pole of the cline, because there are actually many different ways in which a clause can depart from the simple ideal. Thus, compare the following two natural sentences, the first from a conversation and the second from a university textbook:

1. Well, since he got so upset, I just didn’t think we would want to wait for Tina to come back.
   T-unit length: 20
2. This may be part of the reason for the statistical link between schizophrenia and membership in the lower socioeconomic classes.
   T-unit length: 20

The two sentences have identical T-unit length. However, the two are obviously very different in their structural properties, reflected by the number of dependent clauses per T-unit:

1. Well [since he got so upset], I just didn’t think [we would want [to wait for [Tina to come back] ]]]
   main verb: think
   number of dependent clauses per T-unit: 4

2. This may be part of the reason for the statistical link between schizophrenia and membership in the lower socioeconomic classes.
   main verb: be
   number of dependent clauses per T-unit: 0

Which of these two sentences represents greater grammatical complexity? Both of them are complex when contrasted with a simple clause. But these sentences illustrate how misleading it can be to regard complexity as a single unified construct. Rather, the two sentences have fundamentally different grammatical structures, complex in different ways which are likely to cause different kinds of challenges for the language learner.

These two sentences illustrate the problems with both popular T-unit–based measures: the mean length of T-unit and the number of dependent clauses per T-unit (the T-unit complexity ratio):

- Mean length of T-unit: There are many different linguistic devices that can be used to make a long T-unit, including additional dependent clauses, embedded phrases, or even extra adjectives and adverbs. Based on T-unit length, we would conclude that Sentences 1 and 2 are equally complex—disregarding the radically different structures of the two sentences.

---

6 If we count contractions as separate words, the conversation T-unit is one word longer.
Dependent clauses per T-unit: According to this measure, Sentence 2 is not complex at all, whereas Sentence 1 is extremely complex. There are two problems here: first, there are many different kinds of dependent clauses in English, representing different complexities; and second, nonclausal embedding must also be regarded as complex. In fact, we argue below that certain types of nonclausal embedding represent higher orders of complexity than dependent clauses. Minimally, we show below that the complexities of professional written discourse are associated with phrases rather than dependent clauses.

To the extent that T-unit–based measures identify a complexity difference between these two sentences, they produce the wrong result for the study of writing development, categorizing Sentence 1 as more complex than Sentence 2. If we believe T-unit measures, a student who produces Sentence 1 is more developmentally advanced than a student who produces Sentence 2; sentences like 2 should be produced at an earlier developmental stage than sentences like 1.

Intuitively, these conclusions are dissatisfying. Sentence 1 seems relatively commonplace, and there is no difficulty understanding the intended meaning despite the numerous dependent clauses. It is easy to imagine a high school student uttering this sentence in a normal conversation. In contrast, Sentence 2 is more difficult to process, with much information packed into its single clause. It would be very surprising to hear a student produce such a sentence in conversation, and in fact, it would be unusual to encounter a sentence of this type in undergraduate student writing. Thus, it could be argued that Sentence 2 is considerably more difficult than Sentence 1. (Note that the grammatical difficulty of Sentence 2 remains even if we substitute less technical vocabulary, such as disease instead of schizophrenia.) But the important point for our purposes here is that these two sentences represent fundamentally different kinds of complexity.

Such differences led us to question the standard practice of equating complexity with the use of dependent clauses and assuming that such a definition is directly applicable to the assessment of writing development. In contrast, we undertook a bottom-up investigation of complexity, analyzing professional academic written texts to identify the grammatical features that are actually used in complex written discourse. Specifically, we investigated two general research questions:

1. Is extensive subordination typical of professional academic written discourse? That is, would professional academic writing be considered complex by traditional measures of complexity based on T-units?
2. If not, what alternative grammatical devices are used in complex written discourse?

In the following sections, we present the results of a large-scale corpus analysis that addresses these questions. Then, in conclusion, we return to the assessment of complexity in writing development, proposing that these grammatical features discovered through empirical analysis of academic texts provide more appropriate measures of complexity in academic writing development than the traditional measures based on the density of dependent clauses.

EMPIRICAL EVIDENCE: GRAMMATICAL COMPLEXITY IN ACADEMIC WRITING VERSUS CONVERSATION

Observing Complexity in Natural Language Use

Rather than assuming a priori that certain structures are complex, the approach adopted here is based on bottom-up empirical analysis: analyzing academic discourse written by professionals to identify the grammatical features that are most commonly used. Conversational discourse is also analyzed to provide a point of comparison, as the two registers are often contrasted in discussions of grammatical complexity (see, e.g., Hyland, 2002, p. 50). One underlying assumption of this approach is that grammatical structures that are common in conversation do not represent a high degree of production complexity. That is, these structures are regularly and frequently produced by all native speakers of English; they represent normal discourse, and thus there is no evidence from L1 use that they are difficult or complex from a production perspective. In particular, grammatical features that are common in conversation are not appropriate measures for the study of complexity in writing development.

In contrast, grammatical structures that are restricted to academic writing represent the kinds of complexity that must be acquired by advanced students of writing. These structures are highly specialized; they are produced in circumstances that permit careful planning, revising, and editing, and many native speakers of English never acquire the discourse styles that employ these grammatical structures.

The applied objective here is measuring progress in writing development. Clearly, we need an accurate description of the target register—academic writing—in order to determine whether students are progressing toward that target. In particular, it is important to identify complexities that are common in academic writing but rare in conversation, and to distinguish those from complexities that are common in conversation.
Corpus-based analysis is ideally suited for such descriptions of language use. Corpus analyses often produce surprising results, and the present case is no exception: as the following section shows, the kinds of grammatical complexity that are common in academic writing are dramatically different from the grammatical complexities of conversation. And more surprisingly, T-unit measures are much more strongly associated with conversational complexities than the complexities of writing, while a new set of grammatical measures is required to account for the actual complexities of formal written discourse.

Corpus and Grammatical Features Used for the Analysis

As noted above, we employ corpus-based analysis to describe the types of grammatical complexity produced in formal academic writing (written by professionals as opposed to student academic writing). Large-scale corpus analysis is ideally suited to research of this type: a corpus provides a much more representative sample of language than what is typically used in developmental studies; the use of computational techniques enables analysis of these very large text collections, providing results that are generalizable to the target populations; and the use of quantitative analysis allows us to describe the actual extent to which a pattern of use is preferred in one text variety over another. (There are numerous book-length introductions to corpus linguistics; see, e.g., Biber, Conrad, & Reppen, 1998; Hunston, 2002; Kennedy, 1998; McEnery & Wilson 1996; McEnery, Xiao, & Tono, 2006.)

The analysis is based on two large corpora of texts, summarized in Table 1. The subcorpus of academic writing consists of 429 research articles (~3 million words), sampled from four general disciplines: science/medicine, education, social science (psychology), and humanities (history). We collected texts from 11 different academic journals:


| TABLE 1 | Corpus Used for the Analysis
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Register</strong></td>
</tr>
<tr>
<td>Number of texts</td>
</tr>
<tr>
<td>Number of words</td>
</tr>
<tr>
<td>Mean length of text</td>
</tr>
</tbody>
</table>
While there are many subregisters within the broad register of academic writing, we focus on the professional academic writing found in research articles. The four general disciplines listed above were chosen to represent a wide range of research within academia, from hard science to social science to humanities. Within each discipline, articles were chosen randomly and represent a variety of topics.

The conversation subcorpus is taken from Biber et al. (1999; see pp. 24–35). The subcorpus includes 723 text files and ~4.2 million words of American English conversation. These are conversations collected by participants who agreed to carry tape recorders for a 2-week period. The corpus thus represents one of the largest collections of natural face-to-face conversations in existence.

As described above, previous corpus-based studies have documented the different complexities of spoken and written registers (e.g., Biber, 1988, 1992, 2006; Biber et al., 1999). Building on this previous research, the present study focuses on the grammatical devices that are associated with structural complexity, contrasting the distributional patterns of use in conversation versus academic writing. For the analysis, complexity devices are categorized along two structural parameters: grammatical type and grammatical function (see Table 2). Three grammatical types are distinguished: finite dependent clauses, nonfinite dependent clauses, and dependent phrases. These types can serve three major syntactic functions: adverbial, complement, and noun modifier.

As Table 3 shows, there are many specific grammatical subcategories that can be distinguished; some of these are quite frequent in themselves and so are included in our corpus analysis. For example, under finite dependent clauses, there are several subcategories for finite adverbial clause (1A), including causative clauses (because, since), conditional clauses (if), and concessive clauses (although). There are also subcategories of finite clauses as noun modifiers (1C), including relative clauses that begin with that and WH relative clauses (beginning with who, which, whom). Under phrasal structures, there are several specific types of noun modifiers, including adjectives as premodifiers, nouns as premodifiers, and prepositional phrases as postmodifiers (3C).

There are even more distinctions within the general syntactic function of complement clause, distributed across the structural types. For finite complement clauses, there are two major subtypes: that clauses and WH clauses (1B). Similarly, there are two subtypes for nonfinite complement clauses: to clauses and –ing clauses (2B). But the general category of complement clause further includes three specific syntactic functions: (1) complement clauses controlled by a verb, filling a subject or object
slot in the clause; (2) complement clauses controlled by a predicative adjective; (3) complement clauses controlled by a noun. The first syntactic function relates to clause syntax, similar to adverbial clauses, while the third syntactic function relates to the syntax of noun phrases.

### Analytical Procedures

The analysis used an observational research design based on analysis of the academic writing and conversation subcorpora. Each text (i.e., each academic research article or each conversation) was treated as an observation for the purposes of the study: a total of 1,152 observations (see Table 1). The independent variable in the analysis is register: academic writing versus conversation. The dependent variables are the rates of occurrence for each grammatical complexity feature; we analyzed the distributions of 28 specific features. Thus, the first step in the analysis was to compute a normed rate of occurrence for each grammatical feature in each text (e.g., the rate of finite adverbial clauses per 1,000 words). Then we computed mean scores and standard deviations for each grammatical feature in each register, and employed

### TABLE 2

<table>
<thead>
<tr>
<th>Major Grammatical Types and Grammatical Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammatical type</td>
</tr>
<tr>
<td>----------------------------------</td>
</tr>
<tr>
<td>Finite dependent clause</td>
</tr>
<tr>
<td>Complement</td>
</tr>
<tr>
<td>Noun modifier</td>
</tr>
<tr>
<td>Nonfinite dependent clause</td>
</tr>
<tr>
<td>Complement</td>
</tr>
<tr>
<td>Noun modifier</td>
</tr>
<tr>
<td>Dependent phrase (nonclausal)</td>
</tr>
<tr>
<td>Noun modifier</td>
</tr>
</tbody>
</table>
normal inferential statistics to test for the significance and strength of differences between the two registers.

Computer programs were developed to count the occurrences of each complexity feature in each of these texts. Those programs included a grammatical tagger (originally developed for multidimensional analyses

**TABLE 3**

*Specific Structural Distinctions Included in the Analysis*

<table>
<thead>
<tr>
<th>Grammatical type</th>
<th>Grammatical function</th>
<th>Subcategory</th>
</tr>
</thead>
</table>
| 1. Finite dependent clause | 1A. Adverbial | • Causative: She won’t narc on me, because she prides herself on being a gangster.  
• Conditional: Well, if I stay here, I’ll have to leave early in the morning.  
• Concessive: If I don’t put my name, she doesn’t know who wrote it, although she might guess. |
| | 1B. Complement | • Controlled by a verb:  
○ That-clause: I would hope that we can have more control over them.  
○ That-clause (with ZERO): yeah, I think I probably could.  
○ WH-relative: I don’t know how they do it. |
| | 1C. Noun modifier | • Controlled by a noun:  
The fact that no tracer particles were found in or below the tight junction (zonula occludens) indicates that these areas are not a pathway for particles of this size in the toad bladder. |
| 2. Nonfinite dependent clause | 2A. Adverbial | • Purpose: To verify our conclusion that the organic material is arranged as a coating around the silica shell components, thin sections of fixed cells were also examined. |
| | 2B. Complement | • Controlled by a verb:  
○ to-clause: I really want to fix this room up.  
○ ing-clause: I like watching the traffic go by.  
• Controlled by an adjective:  
It was important to obtain customer feedback.  
• Controlled by a noun:  
The project is part of a massive plan to complete the section of road… |
| | 2C. Noun modifier (including –ing clauses and –ed clauses) | • –ing clauses: Transfer tests following over-training indicated individual variability.  
• –ed clauses: The results shown in Tables IV and V add to the picture… |
of register variation; see, e.g., Biber, 1988, 1995) as well as additional programs to identify particular syntactic constructions. The more specific programs incorporated lexicogrammatical information from the *Longman Grammar of Spoken and Written English* (Biber et al., 1999), such as the lists of common nouns that can control a *that* complement clause (e.g., *fact, hypothesis, claim*) or a *to* complement clause (e.g., *proposal, plan, bid*). This allowed us to analyze syntactic features that could not otherwise be identified using automatic techniques. The results of these analyses were checked by hand to ensure accurate identification of the target features.

The syntactic function of prepositional phrases—as adverbial versus postnominal modifier—could not be accurately determined using automatic analysis. The preposition *of* was not problematic, because it almost always functions as a postnominal modifier when it occurs immediately following a noun. However, the prepositions *in, on, with,* and *for* were analyzed by hand. This analysis was based on a subsample of tokens (every fourth occurrence) from a subsample of the corpus (41 academic research articles, and 48 conversations).

As mentioned above, the counts for all linguistic features were converted to a normed rate of occurrence (per 1,000 words) for each text, allowing comparisons across texts of differing lengths (see Biber et al., 1998, pp. 263–264). It was then possible to compute means and standard deviations for each feature in conversation and academic writing. Finally, ANOVA (based on the general linear models procedure in SAS) was used to test whether the differences between registers were statistically significant and to determine the strength of association ($r^2$).
for the differences between the two registers. Register (conversation vs. academic writing) served as the independent variable, whereas the normalized counts of each grammatical feature served as the dependent variables in these tests. The \( r^2 \) scores (computed from the sum-of-squares between groups divided by the total sum-of-squares) provide a measure of the importance or strength of the differences between the two registers.

### Preview of the Corpus Findings: Two Major Parameters of Variation

By this point, it should be obvious why the T-unit has been so favored as a measure of grammatical complexity: it is simple. In contrast, Table 2 lists numerous grammatical types of structural dependencies, each potentially representing a different type of complexity. There are obviously too many different distinctions here to be applied in practice to the assessment of student writing development.

However, when these structures are studied in actual use (see below), it turns out that they pattern along two major parameters:

- **Parameter A: Structural type**
  - Favored in conversation: finite dependent clauses
  - Favored in academic writing: dependent phrases (nonclausal)

- **Parameter B: Syntactic function**
  - Constituents in clauses
  - Constituents in noun phrases

Put simply, the complexity of conversation is clausal, whereas the complexity of academic writing is phrasal. The following subsection presents the findings from our corpus investigation, providing a detailed description of how academic writing differs from conversation with respect to these two parameters.

### Corpus Findings

Tables 4–6 present the results of the statistical comparisons for each complexity feature: Table 4 for finite dependent clause types, Table 5 for nonfinite dependent clause types, and Table 6 for dependent phrase types. As the tables show, most of these complexity features are strongly favored in either conversation or academic writing, but not both. The \( r^2 \) score provides a measure of the importance or strength of the difference between the two registers. For most of these features, the \( r^2 \) value is over 0.3 (i.e., over 30% of the variation in the feature can be predicted by the register difference). However, the \( r^2 \) value for some features is
considerably larger. For example, finite complement clauses controlled by verbs (that clauses and WH clauses) have $r^2$ values over 0.55 (Table 4), while prepositional phrases as noun modifiers have an $r^2$ value around 0.90 (Table 6).

One especially interesting finding here is that many of these complexity features are not common in writing. In fact, these linguistic differences are split between features that are strongly favored in conversation versus features strongly favored in academic writing. Figures 1–3 plot the most frequent features, showing the magnitude of these differences visually.

### Table 4
Statistical Comparisons for Finite Dependent Clause Types

<table>
<thead>
<tr>
<th>Linguistic feature</th>
<th>Conversation mean score</th>
<th>Academic WR mean score</th>
<th>$F$ value</th>
<th>Significance</th>
<th>$r^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finite adverbial clauses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total adverbial clauses</td>
<td>7.1</td>
<td>3.6</td>
<td>603.2</td>
<td>$&lt;0.0001$</td>
<td>0.35</td>
</tr>
<tr>
<td><em>Because</em> clause</td>
<td>2.0</td>
<td>0.6</td>
<td>336.9</td>
<td>$&lt;0.0001$</td>
<td>0.23</td>
</tr>
<tr>
<td><em>If</em> clause</td>
<td>4.0</td>
<td>1.1</td>
<td>749.5</td>
<td>$&lt;0.0001$</td>
<td>0.40</td>
</tr>
<tr>
<td><em>Although</em> clause</td>
<td>0.6</td>
<td>0.05</td>
<td>777.6</td>
<td>$&lt;0.0001$</td>
<td>0.41</td>
</tr>
<tr>
<td>Finite complement clauses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>verb + <em>that</em> clause**</td>
<td>10.8</td>
<td>2.6</td>
<td>2196.7</td>
<td>$&lt;0.0001$</td>
<td>0.66</td>
</tr>
<tr>
<td>verb + WH clause</td>
<td>2.7</td>
<td>0.2</td>
<td>1413.9</td>
<td>$&lt;0.0001$</td>
<td>0.55</td>
</tr>
<tr>
<td>adjective + <em>that</em> clause</td>
<td>0.1</td>
<td>0.3</td>
<td>131.4</td>
<td>$&lt;0.0001$</td>
<td>0.10</td>
</tr>
<tr>
<td>noun + <em>that</em> clause</td>
<td>0.1</td>
<td>0.6</td>
<td>474.1</td>
<td>$&lt;0.0001$</td>
<td>0.29</td>
</tr>
<tr>
<td>Finite noun modifier clauses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>that</em> relative clauses</td>
<td>2.3</td>
<td>2.2</td>
<td>1.8</td>
<td>n.s.</td>
<td>0.29</td>
</tr>
<tr>
<td>WH relative clauses</td>
<td>0.9</td>
<td>3.7</td>
<td>858.1</td>
<td>$&lt;0.0001$</td>
<td>0.43</td>
</tr>
</tbody>
</table>

Note: n.s. = not significant. **Including clauses with a zero complementizer.

### Table 5
Statistical Comparisons for Nonfinite Dependent Clause Types

<table>
<thead>
<tr>
<th>Linguistic feature</th>
<th>Conversation mean score</th>
<th>Academic WR mean score</th>
<th>$F$ value</th>
<th>Significance</th>
<th>$r^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonfinite adverbial clauses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>To</em> adverbial clause</td>
<td>0.08</td>
<td>0.32</td>
<td>172.6</td>
<td>$&lt;0.0001$</td>
<td>0.13</td>
</tr>
<tr>
<td>Nonfinite complement clauses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verb + <em>-ing</em> clause</td>
<td>1.3</td>
<td>0.2</td>
<td>842.5</td>
<td>$&lt;0.0001$</td>
<td>0.42</td>
</tr>
<tr>
<td>Verb + <em>to</em> clause</td>
<td>4.7</td>
<td>3.4</td>
<td>166.6</td>
<td>$&lt;0.0001$</td>
<td>0.13</td>
</tr>
<tr>
<td>Adjective + <em>-ing</em> clause</td>
<td>0.04</td>
<td>0.1</td>
<td>48.6</td>
<td>$&lt;0.0001$</td>
<td>0.04</td>
</tr>
<tr>
<td>Adjective + <em>to</em> clause</td>
<td>0.6</td>
<td>1.3</td>
<td>406.2</td>
<td>$&lt;0.0001$</td>
<td>0.26</td>
</tr>
<tr>
<td>Noun + <em>of</em> + <em>-ing</em> clause</td>
<td>0.05</td>
<td>0.4</td>
<td>310.6</td>
<td>$&lt;0.0001$</td>
<td>0.21</td>
</tr>
<tr>
<td>Noun + <em>to</em> clause</td>
<td>0.9</td>
<td>2.8</td>
<td>856.8</td>
<td>$&lt;0.0001$</td>
<td>0.43</td>
</tr>
<tr>
<td>Nonfinite noun modifier clauses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonfinite relative clause</td>
<td>0.7</td>
<td>4.2</td>
<td>2257.3</td>
<td>$&lt;0.0001$</td>
<td>0.66</td>
</tr>
</tbody>
</table>
Considering the two major parameters introduced in the previous section, the most strongly favored types of structural complexity in conversation are finite dependent clauses functioning as constituents in other clauses. Figure 1 plots the mean scores for finite clauses functioning as adverbials (if and because) and finite clauses functioning as verb complements (that, ZERO, and WH), showing that these clause types are much more common in conversation than in academic writing. Text Sample 1 below illustrates the dense use of these finite clausal structures typical in everyday conversation.

Text Sample 1: Conversation

Finite clauses functioning as adverbials and finite clauses functioning as verb complements are **bold underlined**

<waiting in a car> <very long pause>

Peter: Were Bob and Dorothy up already?

Gayle: Oh yeah they were up. I think **we better wait.** [...] He’s got to have his bacon and egg muffin. We took him to breakfast on Sunday, all he did was complain. <laugh> Of course he gets mad **cause he can’t smoke cause we always take non-smoking.**

Peter: Oh well.

Gayle: See we didn’t know **what we were gonna be doing** and if Karen did go **into labor and we had to leave early or something when we got back** and we wanted to be able to do it. And they wouldn’t take their van

---

**TABLE 6**

Statistical Comparisons for Dependent Phrase Types (Nonclausal)

<table>
<thead>
<tr>
<th>Linguistic feature</th>
<th>Conversation mean score</th>
<th>Academic WR mean score</th>
<th>F value</th>
<th>Significance</th>
<th>$r^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adverbials</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adverbs as adverbials</td>
<td>76.5</td>
<td>28.3</td>
<td>4581.6</td>
<td>&lt;0.0001</td>
<td>0.80</td>
</tr>
<tr>
<td>Prepositional phrases as adverbials*</td>
<td>22.9</td>
<td>31.6</td>
<td>51.06</td>
<td>&lt;0.0001</td>
<td>0.37</td>
</tr>
<tr>
<td>Noun modifiers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attributive adjectives</td>
<td>16.5</td>
<td>57.1</td>
<td>5787.8</td>
<td>&lt;0.0001</td>
<td>0.84</td>
</tr>
<tr>
<td>Nouns as nominal premodifiers</td>
<td>19.0</td>
<td>57.4</td>
<td>1259.2</td>
<td>&lt;0.0001</td>
<td>0.52</td>
</tr>
<tr>
<td>Total prepositional phrases as nominal postmodifiers*</td>
<td>6.3</td>
<td>51.9</td>
<td>1380.1</td>
<td>&lt;0.0001</td>
<td>0.94</td>
</tr>
<tr>
<td>Of as postmodifier</td>
<td>4.6</td>
<td>34.1</td>
<td>9323.4</td>
<td>&lt;0.0001</td>
<td>0.89</td>
</tr>
<tr>
<td>In as postmodifier*</td>
<td>0.5</td>
<td>8.8</td>
<td>152.4</td>
<td>&lt;0.0001</td>
<td>0.64</td>
</tr>
<tr>
<td>On as postmodifier*</td>
<td>0.3</td>
<td>2.5</td>
<td>70.9</td>
<td>&lt;0.0001</td>
<td>0.45</td>
</tr>
<tr>
<td>With as postmodifier*</td>
<td>0.3</td>
<td>2.1</td>
<td>65.1</td>
<td>&lt;0.0001</td>
<td>0.43</td>
</tr>
<tr>
<td>For as postmodifier*</td>
<td>0.7</td>
<td>4.4</td>
<td>107.4</td>
<td>&lt;0.0001</td>
<td>0.55</td>
</tr>
</tbody>
</table>

*Note:* Based on a hand-coded subsample of 89 texts consisting of 41 academic research articles and 48 conversations.
cause Bob wanted to smoke and uh, Ed said he said he’d stop but he can’t smoke in the van.

Peter: Yeah I know.

FIGURE 1. Common finite clause types functioning as clausal constituents.

FIGURE 2. Common dependent phrasal types functioning as constituents in a noun phrase.
Gayle: And Dorothy said Bob's getting terrible with, with the smoking. Uh, he's really getting defiant about it because there are so many restaurants where you can't smoke and he just gets really mad and won't go to them.

[...]
Peter: We went to a party, remember uh, Jim, the veterinarian, that great big guy?
Gayle: Yeah.
Peter: Well they, they had a party. I forget what it was. They had it at a friend’s house. I can’t remember why it wasn’t at their house any way. And they had bought a bottle of Bailey’s because they knew I liked Bailey’s.

[...]
Gayle: Who was it that said last night … how much my dad has loosened up?
Peter: Uh.
Gayle: I can’t remember who it was. One of us kids.

[...]
Peter: Oh. I’ll tell you I think the biggest change in me is since I had my heart surgery.
Gayle: Really?
Peter: My whole outlook is just uh
Gayle: Yeah I guess my, I mean I know my surgery was a good thing but
Peter: <??> It makes you think.
Gayle: Enough to shake me up.
Peter: <??> Yeah. You realize it can happen to you.
Gayle: Yeah.
In contrast, the two parameters are aligned in the opposite way in academic writing: phrasal (nonclausal) structures functioning as constituents in noun phrases (see Figure 2). The use of prepositional phrases as postnominal modifiers is the clearest case of this type and shows the strongest difference: extremely common in academic writing and rare in conversation. Attributive adjectives and nouns as nominal premodifiers are less clear-cut, because they are individual words. However, these are embedded nonclausal constituents in the noun phrase, and they pattern exactly like prepositional phrases as postmodifiers. Text Sample 2 illustrates the dense use of all three nonclausal complexity features typical in academic prose:

Text Sample 2: Academic research article

Prepositional phrases functioning as noun modifiers are **bold underlined**; attributive adjectives are in **italics**; nouns as nominal premodifiers are in **bold italics**

We expected that the use of different transformations would have significant effects on our perceptions of spatial patterns in *kelp holdfast assemblages*. Specifically, we were interested in the qualitative ecological difference in emphasis between changes in composition vs. changes in relative abundance. When analysing presence/absence data, the variability being measured is explicitly the variation in the presence or absence of particular species (or taxa) in different holdfasts at different places: thus compositional change is the essential (and only) feature. This can be contrasted against analyses based on other transformations (or untransformed data) for which variation in relative abundance plays a more important, or even a dominant, role in the analysis.

Finally, Figure 3 plots the use of complexity structures that are mixed or intermediate on the two parameters:

1. finite clause types functioning as a constituent in a noun phrase
2. phrases (nonclausal) functioning as a constituent in a clause
3. nonfinite clause types

There are two things to notice about the distribution of these intermediate features. First, they are generally less frequent in absolute terms than the previous two sets of features. (The exception is phrasal adverbials, especially simple adverbs as adverbials, which are very frequent in conversation.) Second, the differences between conversation and academic writing are less extreme with respect to some of these features than the previous two sets. So, for example, prepositional phrases as adverbials are only moderately more frequent in academic
writing than in conversation \( (r^2 = 0.37) \); verb + to clause constructions are only moderately more frequent in conversation \( (r^2 = 0.13) \); that relative clauses occur with nearly the same frequency in both registers (no significant difference). In general, though, Figure 3 shows that Parameter B (syntactic function) continues to be important, even for the structural categories of finite and nonfinite dependent clause. This is especially the case for dependent clauses functioning as constituents in a noun phrase, which tend to be strongly favored in academic writing (except for that relative clauses). Thus, noun complement clauses (that and to), WH relative clauses, and nonfinite relative clauses are all much more common in academic writing than in conversation, although none of these structures is especially frequent in absolute terms. Text Sample 2, repeated below as Text Sample 3, illustrates these patterns, with three dependent clauses functioning as noun phrase constituents versus only one functioning as a clause constituent.

Text Sample 3: Academic research article

Dependent clauses as noun phrase constituents are **bold underlined**; dependent clauses as clause constituents are in *italics*.

We expected *[that the use of different transformations would have significant effects on our perceptions of spatial patterns in kelp holdfast assemblages]*. Specifically, we were interested in the qualitative ecological difference in emphasis between changes in composition vs. changes in relative abundance. When analysing presence/absence data, the variability *[being measured]* is explicitly the variation in the presence or absence of particular species (or taxa) in different holdfasts at different places: thus compositional change is the essential (and only) feature. This can be contrasted against analyses *[based on other transformations (or untransformed data)]*[ for which variation in relative abundance plays a more important, or even a dominant, role in the analysis].

**DISCUSSION AND CONCLUSION**

The present article has documented the grammatical complexities of academic writing in detail, using conversation as a comparison register. The goals are

- to investigate whether the complexity features traditionally analyzed in studies of writing development are actually characteristic of professional academic writing, and if they are not,
to identify the alternative complexity features that do commonly occur in professional academic writing, providing the basis for new approaches to the study of writing development.

Our findings can be interpreted in a strong or weak manner. The weaker interpretation of the findings is not controversial: The kinds of complexity common in academic writing are fundamentally different from the kinds of complexity common in conversation. Thus, complexity is not a single unified construct, and it is therefore not reasonable to suppose that any single measure will adequately represent this construct.

In particular, the T-unit is a measure designed mostly to capture the extent to which a writer uses dependent clauses. As a result, this measure misses out on the most important kinds of complexity devices in academic writing: nonclausal features embedded in noun phrases. Thus, we need additional measures to capture development toward the kinds of grammatical complexities that are most important in academic writing.

The strong interpretation begins with the developmental progression implied by the comparison of conversation and academic writing: Conversation is acquired first; the grammar of writing is acquired later, and not always successfully. Grammatical structures that are readily acquired (at relatively early stages) and frequently produced in conversation by all native speakers of a language are obviously not difficult; therefore these structures do not represent a high degree of production complexity. In contrast, many types of complex phrasal embedding are produced in only the more specialized circumstances of formal writing. These styles of discourse are not acquired naturally, and many native speakers of English rarely (or never) produce language of this type. Further, when these stages of acquisition do occur, they are late, typically in adulthood. Considering all these factors, it is reasonable to hypothesize that these grammatical structures represent a considerably higher degree of production complexity than the conversational complexity features.

Based on the observed developmental patterns for L1 learners, we can hypothesize a similar series of developmental stages for L2 learners of English, mirroring the progression from conversational competence to competence in academic writing. Of course, some L2 learners never acquire conversational skills, being taught written rather than spoken English. However, even for those students, competence in English academic writing is developed late, and thus the complexity features of academic writing will be acquired in later developmental stages.

The two parameters of variation described in the last section are the major determining factors for the progression. Thus, the stages generally progress from finite dependent clauses functioning as
constituents in other clauses, through intermediate stages of nonfinite dependent clauses and phrases functioning as constituents in other clauses, and finally to the last stage requiring dense use of phrasal (nonclausal) dependent structures that function as constituents in noun phrases.

In addition, we have considered lexicogrammatical factors in positing these developmental stages, because dependent clause structures in conversation tend to occur with just a few controlling words, while

### TABLE 7

<table>
<thead>
<tr>
<th>Stage</th>
<th>Grammatical structure(s)</th>
<th>Example(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Finite complement clauses (<em>that and WH</em>) controlled by extremely common verbs (e.g., think, know, say)</td>
<td>1a we never quite <strong>know</strong> what to make of him (conv)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1b just <strong>think</strong> that he didn’t pay attention (conv)</td>
</tr>
<tr>
<td>2</td>
<td>Finite complement clauses controlled by a wider set of verbs Finite adverbial clauses</td>
<td>2a I’d <strong>forgotten</strong> that he had just testified on that one (conv)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2b <strong>If you’re sitting next to me and you want ninety degrees, and I want sixty degrees, we’re just gonna be battling each other…</strong> (conv)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2c I’m assuming I gained weight because things are a little tighter than they used to be (conv)</td>
</tr>
<tr>
<td></td>
<td>Nonfinite complement clauses, controlled by common verbs (especially want)</td>
<td>2d I don’t <strong>want</strong> to fight with them about it (conv)</td>
</tr>
<tr>
<td></td>
<td>Phrasal embedding in the clause: adverbs as adverbials</td>
<td>2e I <strong>hate</strong> watching the people interact (conv)</td>
</tr>
<tr>
<td></td>
<td>Simple phrasal embedding in the noun phrase: attributive adjectives</td>
<td>2f <strong>He’s so confused anyway</strong> (conv)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2g It certainly has a nice flavor (conv)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2h <strong>Tom Jones is apparently a real name</strong> (conv)</td>
</tr>
<tr>
<td>3</td>
<td>Phrasal embedding in the clause: prepositional phrases as adverbials</td>
<td>3a He seems to have been hit on the head (fict)</td>
</tr>
<tr>
<td></td>
<td>Finite complement clauses controlled by adjectives</td>
<td>3b It seemed quite <strong>clear</strong> that no one was at home (fict)</td>
</tr>
<tr>
<td></td>
<td>Nonfinite complement clauses controlled by a wider set of verbs</td>
<td>3c I was <strong>sure</strong> that I could smooth over our little misunderstanding (fict)</td>
</tr>
<tr>
<td></td>
<td><em>That</em>* relative clauses, especially with animate head nouns</td>
<td>3d The snow began to fall again (fict)</td>
</tr>
<tr>
<td></td>
<td>Simple phrasal embedding in the noun phrase: nouns as premodifiers</td>
<td>3e <strong>…the guy that made that call</strong> (fict)</td>
</tr>
<tr>
<td></td>
<td>Possessive nouns as premodifiers</td>
<td>3f <strong>…some really obscure cable channel</strong> (fict)</td>
</tr>
<tr>
<td></td>
<td><em>Of</em>** phrases as postmodifiers</td>
<td>3g <strong>Tobie’s voice</strong> (fict)</td>
</tr>
<tr>
<td></td>
<td>Simple PPs as postmodifiers, especially with prepositions other than <em>of</em> when they have concrete/locative meanings</td>
<td>3h <strong>editor</strong> of the food section (fict)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3i <strong>house</strong> in the suburbs (fict)</td>
</tr>
</tbody>
</table>
structures in academic writing occur with a much wider range of controlling words. For example, \(~75\%\) of all *that* complement clauses in conversation occur with only four verbs: *think, say, know, guess* (see Biber et al., 1999, p. 668). *To* clauses are generally not frequent in conversation, but the combination *want + to clause* is extremely common (see Biber et al., 1999, p. 711). We thus hypothesize that these relatively fixed lexicogrammatical combinations are acquired at an earlier stage than full control of the target syntactic structure. Table 7 presents a specific hypothesis of what these developmental stages might consist of.

Of course, these hypothesized developmental stages need to be investigated by empirical studies of L2 language development. However, this framework has the advantage of being based on the empirical study of actual language production, rather than being posited purely on intuitive or theoretical grounds. This is an important point: the complexity measures used in current research are *not* derived from

---

**Table 7**

<table>
<thead>
<tr>
<th>Stage</th>
<th>Grammatical structure(s)</th>
<th>Example(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 4</td>
<td>Nonfinite complement clauses controlled by adjectives</td>
<td>4a These will not be easy to obtain (acad)</td>
</tr>
<tr>
<td></td>
<td>Extraposed complement clauses</td>
<td>4b It is clear that much remains to be learned... (acad)</td>
</tr>
<tr>
<td></td>
<td>Nonfinite relative clauses</td>
<td>4c In that case it is useful to phrase sustainability in terms of... (acad)</td>
</tr>
<tr>
<td></td>
<td>More phrasal embedding in the NP = attributive adjectives, nouns as premodifiers</td>
<td>4d ...the method used here should suffice... (acad)</td>
</tr>
<tr>
<td></td>
<td>Simple PPs as postmodifiers, especially with prepositions other than <em>of</em> when they have abstract meanings</td>
<td>4e Studies employing electrophysiological measures (acad)</td>
</tr>
<tr>
<td></td>
<td>Preposition + nonfinite complement clause</td>
<td>4f The prevalence of airway obstruction and self-reported disease status (acad)</td>
</tr>
<tr>
<td></td>
<td>Complement clauses controlled by nouns</td>
<td>4g Positive propagule size effects have been demonstrated for both plant and animal systems</td>
</tr>
<tr>
<td></td>
<td>Appositive noun phrases</td>
<td>4h with half of the subjects in each age/instructional condition receiving each form (acad)</td>
</tr>
<tr>
<td></td>
<td>Extensive phrasal embedding in the NP: multiple prepositional phrases as postmodifiers, with levels of embedding</td>
<td>4i The specific growth rate at small population sizes...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5a The idea of using a Monte Carlo approach (acad)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5b The hypothesis that female body weight was more variable (acad)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5c The CTBS (the fourth edition of the test) was administered in 1997–1998 (acad)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5d The [presence of layered [[structures] at the [[[borderline]]] of cell territories]] (acad)</td>
</tr>
</tbody>
</table>
consideration of student writing. Rather, they have been selected on an a priori basis and then applied to the analysis of student writing. We are instead proposing here a principled basis for the selection of complexity measures that can be applied to the evaluation of student writing.

In sum, measures of subordination capture only one kind of grammatical complexity, while the T-unit confounds a wide range of different devices that can be used to create longer structures. Both types of complexity measures miss out on the most important kinds of complexity devices in academic writing: nonclausal features embedded in noun phrases. Whether the strong interpretation of our findings stands or not, it is clear that the grammatical complexities of writing are different from those of speech, that currently used measures reflect the latter rather than the former, and that we therefore need additional measures to capture development toward the complex styles of academic writing.

THE AUTHORS

Douglas Biber is Regents’ Professor of English (applied linguistics) at Northern Arizona University, in the United States. His research efforts have focused on corpus linguistics, English grammar, and register variation.

Bethany Gray is a PhD candidate in the applied linguistics program at Northern Arizona University, in the United States. Her interests include corpus linguistics, register variation, and disciplinary writing.

Kornwipa Poonpon is a lecturer of English at Khon Kaen University, in Thailand. She specializes in second language assessment, corpus linguistics, and teaching English as a second/foreign language.

REFERENCES


