Article

Douglas Biber* and Tove Larsson

Accounting for the entire system of complexity features: evidence for general oral versus literate grammatical complexity dimensions

https://doi.org/10.1515/cllt-2025-0017
Received February 13, 2025; accepted May 23, 2025; published online June 11, 2025

Abstract: Two recent studies (Biber, Douglas, Larsson Tove & Gregory R. Hancock. 2024a. The linguistic organization of grammatical text complexity: Comparing the empirical adequacy of theory-based models. *Corpus Linguistics and Linguistic Theory* 20(2). 347–373; Biber, Douglas, Larsson Tove & Gregory R. Hancock. 2024b. Dimensions of text complexity in the spoken and written modes: A comparison of theory-based models. *Journal of English Linguistics* 52(1). 65–94) evaluate how well theory-based models account for the distributions of complexity features across spoken and written texts. Those studies provide strong evidence for two groupings of complexity features: phrases functioning syntactically as noun modifiers, and finite dependent clauses functioning as clause-level constituents. At the same time, those studies fail to identify systematic patterns of covariation for the other complexity features (e.g., phrases functioning as clause-level constituents). The present study picks up where those previous studies left off, exploring the possibility that complexity features pattern together in systematic ways at the *register* level, even though they have less strong patterns of covariation across individual texts. The results show that: 1) all 25 features can be grouped into one of two groupings, referred to as the "oral" and "literate" complexity dimensions, and 2) those two dimensions have a strong complementary relation to one another. These general patterns are described and interpreted relative to the particular features grouped into each dimension, the register distributions associated with each dimension, and the extent to which these register-level patterns are found at the text level.

***Corresponding author: Douglas Biber**, Northern Arizona University, Flagstaff, 86001-5766, USA, E-mail: douglas.biber@nau.edu
Tove Larsson, Northern Arizona University, Flagstaff, 86001-5766, USA. https://orcid.org/0000-0002-0489-2697

Keywords: grammatical complexity; oral complexity; literate complexity; complexity dimensions; spoken and written register variation

1 Introduction

Over the last decade, there has been an explosion of studies that investigate the grammatical complexity of texts. These studies are especially prevalent in applied linguistics journals, where they usually focus on the complexity of texts written by language learners (see, e.g., Bulté and Housen 2018; Casal and Lee 2019; Kushik and Huhta 2020; Lu 2017; see also the survey of 20+ recent publications in Biber et al. 2025a, 2025b, pp. 147–148). In addition, several recent studies analyze the complexities of texts from different spoken and written registers (see, e.g., the chapters in Biber et al. 2022).

One major theme that runs across many of these studies is the question of the best way to analyze or measure grammatical complexity in texts, relating to underlying questions concerning the nature of the construct itself. There is widespread agreement that grammatical complexity, referring to structural characteristics of grammatical constructions, needs to be distinguished from cognitive difficulty (see the discussion in Bulte et al. 2025). There is also general agreement on the strictly linguistic definition of grammatical complexity as the addition of optional structural elements to a "simple" phrase or a "simple" clause (where "simple" phrases/clauses are defined as structures that include only obligatory elements plus accompanying function words). Traditionally, grammarians have focused on dependent clauses as the most important manifestation of phrase/clause complexity (see, e.g., Huddleston 1984: 378; Willis 2003: 192; Purpura 2004: 91; Carter and McCarthy 2006: 489).

However, beyond that general agreement, there has been considerable debate over the particular measures and variables to use for the study of grammatical complexity in actual language use (sometimes referred to as "text complexity"; see Szmrecsanyi 2015: 347–349). Two major approaches have been especially prevalent: the "omnibus" approach, which relies on omnibus measures that are intended to capture the entire construct of complexity, versus the "Register-Functional" (RF) approach, which distinguishes among specific grammatical complexity features that have different structural characteristics, serving different syntactic functions, with different patterns of register variation. Previous publications (e.g., Biber et al. 2020; Biber et al. 2025b) have argued against the omnibus approach on the grounds that it confounds and/or disregards grammatical distinctions. Specifically, the omnibus approach collapses consideration of multiple grammatical structures, and it completely disregards consideration of syntactic function. In contrast, distinguishing among the range of grammatical structures and their different syntactic functions are foundational to analyses carried out in the RF approach.

These different methodological frameworks reflect different underlying conceptualizations of the construct of grammatical complexity in language use. Our goal here is not to rehash the strengths and weaknesses of each of these approaches. Rather, we are assuming the validity of the research findings from two recent studies carried out by Biber, Larsson, and Hancock. In particular, we assume the importance of two major groupings of complexity features identified in those studies: phrases functioning syntactically as noun modifiers, and finite dependent clauses functioning syntactically as clause-level constituents. The present study picks up where those previous studies left off, addressing the major unresolved question from that body of research: are there systematic patterns of covariation among the other complexity features in English? That is, by employing alternative methods, are we able to account for the language-use patterning of the entire system of complexity features? The background to this question is explained in more detail in Sections 1.1 and 1.2 below.

1.1 Summary of the Biber, Larsson, and Hancock studies

Two recent studies carried out by Biber et al. (2024a, 2024b – referred to as the *BLH studies* here) statistically compare the adequacy of theory-based models based on the predictions made by the omnibus versus RF methodological approaches. These studies are based on the assumption that linguistic features that function as part of the same underlying construct will co-occur regularly in texts and covary in systematic ways across texts.

In practice, the omnibus approach predicts that all complexity features should co-occur in texts as a unified group (because analyses in this approach do not distinguish systematically among specific grammatical structures or syntactic functions). In contrast, the RF approach – which distinguishes among specific structural types and syntactic functions – predicts that groupings of features that share structural/syntactic characteristics should co-occur regularly in texts.

To test those competing hypotheses, the BLH studies computed the rates of occurrence for each complexity feature in each text and then employed a confirmatory approach from the structural equation modeling framework (based on Pearson correlations and the Standardized Root Mean Squared Residual measure) to capture the extent to which each hypothesized model (i.e., the hypothesized groupings of complexity features) is represented by features that actually co-occur regularly in texts; that is, to what extent the hypothesized models fit the data.

The model associated with the RF approach adopts the inventory of grammatical complexity features cataloged in descriptive grammars of English. That is, from a formal perspective, grammatical complexity in English is itself a highly complex

system, including many different types of grammatical structures serving many different syntactic functions. Those grammatical structures include different types of phrases (e.g., noun phrases, adjective phrases, adverb phrases), different types of finite dependent clauses (e.g., *that*-clauses, *WH*-clauses, finite adverbial clauses), and different types of nonfinite dependent clauses (e.g., *to*-clauses, *ing*-clauses). In addition, grammatical structures serve different syntactic functions, such as modifying a head noun, modifying an adjective, complementing a verb, or as a clause-level adverbial.

Importantly, the same type of grammatical structure can be used to serve different syntactic functions. For example, a finite dependent *WH*-clause can function as a noun modifier (1), a clause-level adverbial (2), or a verb complement (3):

(1) *That's a conclusion* ***which has no supporting evidence.*** [relative clause: noun modifier]
(2) *We take them into account* ***when we draw conclusions.*** [clause-level adverbial]
(3) *I don't know* ***how they do it.*** [verb complement]

And conversely, a single syntactic function can be realized as different grammatical structures. For example, the syntactic function of modifying a noun phrase can be realized by a phrase (4), a nonfinite dependent clause (5), or a finite dependent clause (6):

(4) *The scores* ***for male and female students*** *were combined.* [prepositional phrase]
(5) *This is a phrase* ***used in the recruitment industry.*** [nonfinite clause]
(6) *… the experimental error* ***that could result from using cloze tests.*** [finite clause]

The full set of structures and syntactic functions are explained and illustrated in any descriptive grammar (e.g., Biber et al. 1999/2021; Huddleston and Pullum 2002; Quirk et al. 1985).

Table 1 provides a schematic representation of the major complexity features in the grammatical system of English. The columns in Table 1 represent three major structural types of complexity features, while the rows represent three major syntactic functions that complexity features can serve. Each cell in the table shows the specific features that have those characteristics (Appendix provides more details and examples of these complexity features).

The BLH studies provide strong empirical evidence showing that it is essential to distinguish among the structural types of complexity features (e.g., finite dependent clauses, nonfinite dependent clauses, dependent phrases) as well as their syntactic functions (e.g., noun modifiers, other phrase-level constituents, clause-level

Table 1: Phrasal/clausal complexity features in the grammatical system of English, organized according to major structural type and syntactic function.

Syntactic function	Structural type		
	(A) Finite dependent clauses	**(B) Nonfinite dependent clauses**	**(C) Dependent phrases**
1. Noun phrase modifiers	1A Finite relative clauses [FiniteRel] Noun + *that* complement clauses [Noun + THAT]	1B Passive (*-ed*) nonfinite relative clauses [EDRel] *-ing* nonfinite relative clauses [INGRel] *To* relative clauses [TORel] Noun + *to* complement clauses [Noun + TO]	1C Attributive adjectives [Adj + N] Premodifying nouns [N+ N] *Of* genitive phrases [N+ OF] Other postnominal prepositional phrases [N+ Prep]
2. Other phrase complements/modifiers	2A Adjective + *that* complement clauses [Adj + THAT] Preposition + *WH* complement clauses [Prep + WH]	2B Adjective + *to* complement clauses [Adj + TO] Preposition + *-ing* complement clauses [Prep + ING]	2C Adjective + prepositional phrase [Adj + Prep] Adverbs as adjective/adverb modifier [AdvMod]
3. Clause constituents	3A Finite adverbial clauses [FiniteAdvlCls] Verb + *that* complement clauses [Verb + THAT] Verb + *WH* complement clauses [Verb + WH]	3B Verb + *to* complement clauses [Verb + TO] Verb + *-ing* complement clauses [Verb + ING] *To* adverbial clauses [TOAdvl] *-ing* adverbial clauses [INGAdvl]	3C Clause-level adverbs [AdvAdvl] Clause-level prepositional phrases [PrepAdvl]

constituents). Grouping complexity features according to their syntactic functions, especially when combined with structural types, provides the best fit to the observed patterns of covariation.

Two specific subgroups of complexity features receive especially strong empirical support from those analyses: Cell 1C and Cell 3A.

- Cell 1C: dependent phrases functioning syntactically as noun modifiers (highlighted in yellow in Table 1)
  - attributive adjectives, premodifying nouns, *of*-genitive phrases, other prepositional phrases postmodifying a noun

- Cell 3A: finite dependent clauses functioning syntactically as clause-level constituents (highlighted in blue in Table 1)
  - finite adverbial clauses, verb + *that* complement clauses, verb + *wh* complement clauses

These two groups can be regarded as stereotypical phrasal complexity versus stereotypical clausal complexity: the 1C group of features is both structurally and syntactically phrasal, while the 3A group is both structurally and syntactically clausal. Previous research (e.g., Biber et al. 2024b and the chapters in Biber et al. 2022) has shown that these two complexity groupings are related in terms of their discourse functions and register distributions. Finite dependent clauses as clause-level constituents (Group 3A) are especially frequent in spoken registers, functioning to express personal stance meanings or to situate discourse relative to situated (adverbial) meanings. In contrast, dependent phrases as noun modifiers (Group 1C) are especially frequent in informational written registers, functioning to compress maximal amounts of information into relatively few words.

In addition, the BLH studies provided strong evidence that these two subgroups have a systematic relation with one another, usually occurring in complementary distribution. That is, when a text frequently employs finite dependent clauses functioning as clause-level constituents, that same text will usually disfavor the use of phrases functioning as noun modifiers – and vice versa. Thus, the two groups can be interpreted as two complementary dimensions that together constitute a higher-order parameter of variation, opposing stereotypical-phrasal complexity features (i.e., phrases functioning syntactically as noun modifiers) versus stereotypical-clausal complexity features (i.e., finite dependent clauses functioning as clause-level constituents). That is, at some level, these are two fundamentally different types of grammatical complexity that compete with one another, so that speakers/writers tend to employ one set of features or the other.

1.2 Unresolved questions from the BLH studies

Beyond the strong systematic patterns of variation summarized in the last subsection, the BLH studies also showed that there is much that we still do not understand about the functional organization of complexity features. The major unresolved mystery concerns the remaining seven cells in Table 1: for the most part, the features within each of those cells do not tend to have strong patterns of covariation across

texts.[1] That is, the BLH analyses provide little evidence that these other complexity features work together in texts as groupings that are organized either by their structures or by their syntactic functions. Thus, we are left with the possibility that the complexity system of English is organized functionally (i.e., in actual language use) as one very strong higher-order parameter of variation opposing stereotypical-phrasal complexity versus stereotypical-clausal complexity, coupled with a large number of other complexity features that "have their own peculiar distributions, apparently motivated by specialized discourse functions" (Biber et al. 2024a: 371).

While this is one possible explanation of the findings in the BLH papers, it is not a very satisfying one. It would suggest that grammatical complexity – apart from the 1C versus 3A bipolar opposition – is not really a construct from the perspective of language use. That is, apart from the strong stereotypical phrasal-clausal opposition, this explanation would simply posit a collection of 17 other structural-syntactic features that do not pattern together systematically in texts and therefore do not reflect any kind of underlying discourse construct.

There are, however, other possible explanations for the BLH findings. One alternative explanation is that there could be functional groupings of complexity features that do not correspond to either the structural distinctions or the syntactic distinctions shown in Table 1. That is, the BLH analyses tested only the hypothesis that the features within each cell should positively correlate with one another. But it might be the case that features across cells will have positive correlations, indicating that they co-occur in texts because they share discourse functions regardless of their structural/syntactic characteristics.

We explore this possibility in Section 3.1 below. It turns out, though, that this possibility proved to be largely unproductive. That is, the results presented in Section 3.1 below show that most of the complexity features in Table 1 – apart from the 1C and 3A features – simply do not strongly correlate with other complexity features across texts. That finding led us to consider the possibility that complexity features have a higher-level functional organization: that they are organized into groups of features that are associated with registers, even though those features do not systematically co-occur in texts. That is, the system of complexity features might be organized as groups based on their similar patterns of variation across registers, even in the case when those features do not systematically co-occur within the same texts.

The corpus-based analyses of complexity features in the *Grammar of Spoken and Written English* (GSWE; Biber et al. 2021) provide some support for this possibility. The quantitative analyses in the GSWE simply report overall rates of occurrence for

1 The two features in Cell 1A – finite relative clauses and finite *that* noun complement clauses – are exceptional, in that the analyses in BLH 2024 (Table 5) show that those two features do tend to co-occur to some extent in texts.

each register, rather than a true mean score based on analysis of all texts within a register. However, comparing the rates for multiple features across registers indicates that sets of features pattern in similar ways across registers. That is, we do not know from these analyses whether features actually co-occur in the same texts, but we can identify sets of features that follow similar patterns of variation across registers.

For example, Figures 1–3 summarize corpus-based findings relating to the register distribution of five complexity features (based on GSWE, Figure 8.7, 8.13, and 8.23). These results are presented in three separate figures, because the absolute frequencies of the features are strikingly different: attributive adjectives and prepositional phrases as noun phrase modifiers are extremely common (Figure 1); *ed* clauses and *ing* clauses as noun phrase modifiers are moderately common (Figure 2); and *that* noun complement clauses are relatively rare (Figure 3). However, comparing the trends across the three figures shows that all of these features vary in a similar way across registers: most common in academic writing, intermediate frequencies in newspaper writing and fiction, and least common in conversation.

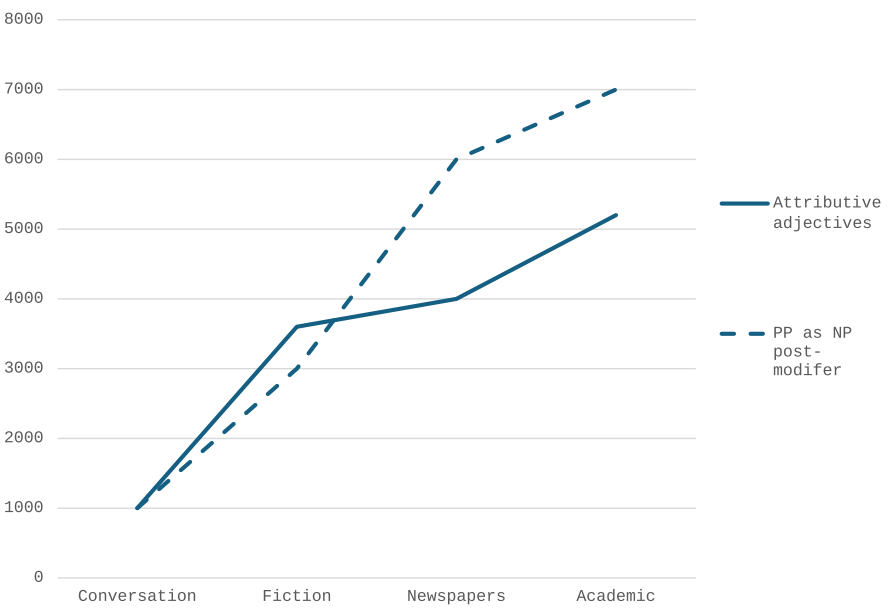

Figure 1: Variation across registers for 1C features: attributive adjectives and prepositional phrases as noun modifiers (per 100,000 words) [based on GSWE].

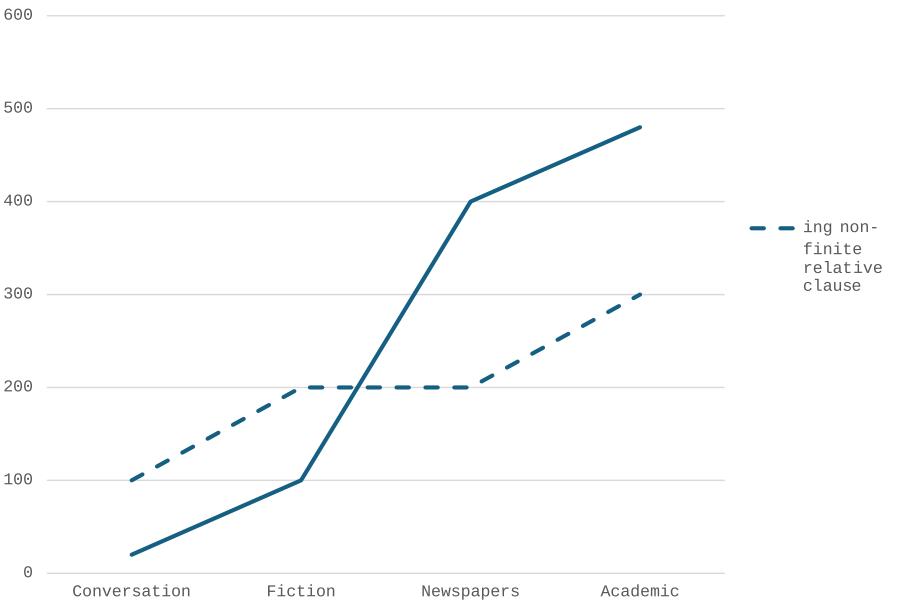

Figure 2: Variation across registers for 1B features: nonfinite *ing*-clauses and *ed*-clauses as noun modifiers (per 100,000 words) [based on GSWE].

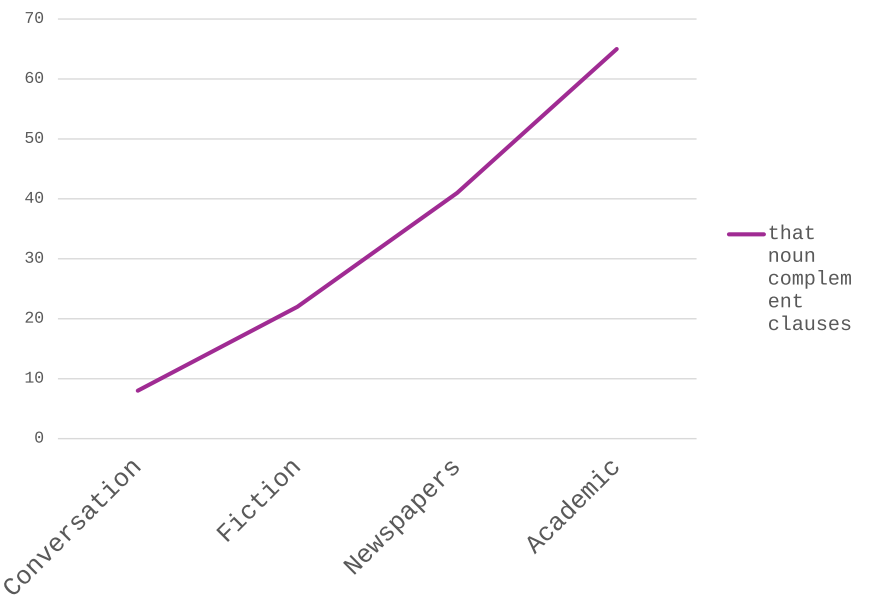

Figure 3: Variation across registers for a 1A feature: finite *that*-clauses as noun complements (per 100,000 words) [based on GSWE].

These five complexity features are from three different cells in Table 1. All five features function syntactically as noun phrase modifiers (or noun complements), but they differ in their structural types: attributive adjectives and prepositional phrases (Figure 1) are phrasal modifiers (cell 1C); *ed* clauses and *ing* clauses (Figure 2) are nonfinite dependent clauses (cell 1B); and *that* noun complement clauses (Figure 3) are a type of finite dependent clause (cell 1A). As noted above, the five features also differ in their overall rates of occurrence: the phrasal features are extremely common (Figure 1; thousands of occurrences in 100,000 words of text); the nonfinite clause features are moderately common (Figure 2; hundreds of occurrences in 100,000 words of text); while *that* noun complement clauses (Figure 3) are relatively rare. But despite those differences, all five of these features follow a similar pattern across registers: comparatively rare in spoken conversation, and comparatively frequent in newspaper writing and academic prose, with the rates of occurrence for fiction usually being intermediate between those two extremes. Research findings like these suggest the possibility that some complexity features are related functionally at the register level, even if they do not regularly co-occur at the text level.

This is the major possibility that we explore in the present paper. Our overall goal is to investigate whether a register-level perspective can help to explain the two major contrasting findings documented in the BLH studies:

– the very strong text-level correlations among all stereotypically phrasal complexity features (Cell 1C), and among all stereotypically clausal complexity features (Cell 3A) – coupled with the very strong inverse text-level correlations between the 1C and the 3A sets of features.

  as well as

– the weak text-level correlations among the complexity features within each of the other seven cells of Table 1.

Specifically, we address the major unresolved question from the BLH studies: are there systematic patterns of covariation among the other complexity features in English? That is, by employing alternative methods, are we able to account for the language-use patterning of the entire system of complexity features?

To achieve this goal, we begin by analyzing text-level correlations among all 25 of the complexity features listed in Table 1, exploring the possibility that complexity features co-occur in texts in groupings that are not represented by either their structural types or syntactic functions. Then, we extend that analysis to consider register-level correlations among the 25 complexity features, exploring the possibility that complexity features covary across registers even if they do not regularly co-occur within texts (similar to the patterns of variation from the *GSWE* illustrated above).

2 Corpus and methods

The corpora and linguistic features analyzed for the study are identical to those analyzed in the BLH studies, based on quantitative-linguistic analyses of the 25 phrasal/clausal complexity features listed in Table 1 in a multiregister corpus. Our goal in the design of the corpus, summarized in Table 2, was to cover a wide range of register variation within each mode, while at the same time generally matching spoken and written registers for their communicative purposes, level of assumed expertise, and interactivity, to avoid confounding the influence of mode with other situational factors. All quantitative-linguistic analyses are based on the normed rates of occurrence (per 1,000 words) for each feature in each text. Additional details of the corpus and quantitative-linguistic analyses are given in BLH (2024a: 14–15).

The goal of the analyses here is to further explore the functional organization of complexity features. We continue to begin with the assumption that linguistic features that function in similar ways will co-occur in language use. The BLH studies analyzed the extent to which those co-occurrence patterns exist within and across texts. We retain this approach for our first analytical step in the present paper (Section 3.1 below). Then, in the second analytical step, we extend this assumption to the register level, exploring the extent to which co-occurrence patterns exist within and across registers (see Section 3.2 below). For all analyses, patterns of co-occurrence are captured through simple Pearson correlations, based on the assumption that complexity features that co-occur and "function together" should have relatively large positive correlations with one another.

Table 2: Breakdown of the corpus across texts and registers.

	# of texts	# of words
Spoken registers		
Conversation	529	231,000
Classroom teaching	145	1,038,000
Formal university lectures	94	593,000
Written registers		
Opinion blogs	177	263,000
Fiction	89	2,226,000
Newspaper articles	100	122,000
University textbooks	56	488,000
Academic research articles	342	2,367,000
Total	1,532	c. 7,300,000

As described above, the present study is motivated by the fact that the BLH studies failed to find strong patterns of covariation for the complexity features in seven of the nine cells of Table 1 (Cells 1A, 2A, 1B, 2B, 3B, 2C, 3C). This fact might simply mean that *complexity* is a structural/syntactic construct, but not a well-defined functional construct. It is also possible, though, that the broader inventory of complexity features pattern together as a functional construct in ways that were not captured by the BLH studies. We explore two of those alternative analytical perspectives here.

The first analysis is based on the same correlation matrix that we used in BLH (2024a). In that study, we began by computing Pearson correlations to measure the extent to which two features co-occur in texts. However, the subsequent analyses in the BLH study took a theory-driven approach, testing the adequacy of the theoretical groupings shown in Table 1 (based on the requirement that each complexity feature in a hypothesized grouping should have a correlation >+0.2 with all other features in that same group). In contrast, the first analysis in the present study takes a step back, considering this same correlation matrix from an exploratory perspective to look for any meaningful correlations among complexity features (i.e., regardless of their structural/syntactic characteristics). That is, we adopt a purely bottom-up exploratory approach, asking simply if any complexity features (apart from the correlations among the 1C and 3A features) co-occur in texts, regardless of their structural type and regardless of their syntactic function.

In contrast, the second analytical step systematically explores the extent to which complexity features can be grouped into functional sets based on their shared patterns of covariation across registers. For that analysis, we begin by computing the mean score of each complexity feature in each of the eight registers listed in Table 2. Then, we compute a new correlation matrix (using Pearson correlations) for all 25 complexity features, again asking if any of the complexity features covary. This approach differs from the first analysis in that we are here investigating the extent to which sets of complexity features covary in the same way across registers, even if those features do not regularly co-occur in the same texts. We recognize the need for caution when interpreting the correlational results from this approach: correlations based on the mean scores of register categories are likely to be larger simply because we are disregarding the variation across texts within registers. However, our hope was that this analysis might uncover an organization to the complexity system that would not be apparent otherwise.[2]

2 This approach represents a radical methodological departure from previous studies carried out in the Text-Linguistic Approach to register variation; we discuss the theoretical implications of that departure in Section 5 below.

3 Results

3.1 Text-level correlations

Figure 4 presents a correlation matrix showing the extent to which each pair of complexity features systematically covaries across texts. We present the results in the form of a heat map, which uses the intensity of colors to represent the magnitude of the correlation. Blue represents positive correlations, and red represents inverse (or negative) correlations. Thus, cells with the darkest blue represent the strongest positive correlations (like the correlation of +0.8 between N + OF and Adj + N), while cells with the darkest red represent the strongest inverse correlations (like the correlation of −0.7 between AdvAdvl and Adj + N). To aid in the interpretation of results, we have organized the variables in Figure 4 in the same order as Table 1, proceeding from 1A to 1B to 1C to 2A, etc.

The results shown in Figure 4 provide the details (i.e., the pairwise correlations) for the statistical findings regarding Cells 1C and 3A in BLH (2024a). Thus, each of the features in Cell 1C (Adj + N, N + N, N + OF, N + Prep) have strong positive correlations with the all the other features in this cell (ranging from +0.54 to +0.80). Similarly, each of the features in Cell 3A (FiniteAdvlCls, Verb + THAT, Verb + WH) have strong or moderately strong positive correlations with all the other features in that cell (ranging from +0.47 to +0.38). And finally, each of the features in Cell 1C have strong or moderately strong inverse (negative) correlations with the each of the features in Cell 3A (ranging from −0.65 to −0.35).

Beyond that pattern, there are relatively few pairwise comparisons in Figure 4 that show even moderately strong correlations between complexity features. In addition, it is noteworthy that those other correlations do not indicate the existence of additional groupings beyond 1C and 3A. Rather, to the extent that there are other moderately strong correlations, they occur between an additional feature and either the 1C grouping or the 3A grouping (i.e., rather than forming the nucleus for a third major grouping). Three features are especially noteworthy in this regard: EDRel, Prep + ING, and AdvAdvl. EDRel and Prep + ING both have strong or moderately strong positive correlations with all the features in Cell 1C (ranging from +0.43 to +0.55), but no strong correlations with any other complexity feature. These features also have inverse correlations with all 3A features (ranging from −0.16 to −0.31). AdvAdvl has strong or moderately strong positive correlations with all the features in Cell 3A (ranging from +0.44 to +0.60), but no strong correlations with any other complexity feature. And this feature has very strong negative correlations with all the features in Cell 1C (ranging from −0.85 to −0.94). Thus, the noteworthy finding here is that the 1C versus 3A opposition (documented in BLH 2024a) actually includes

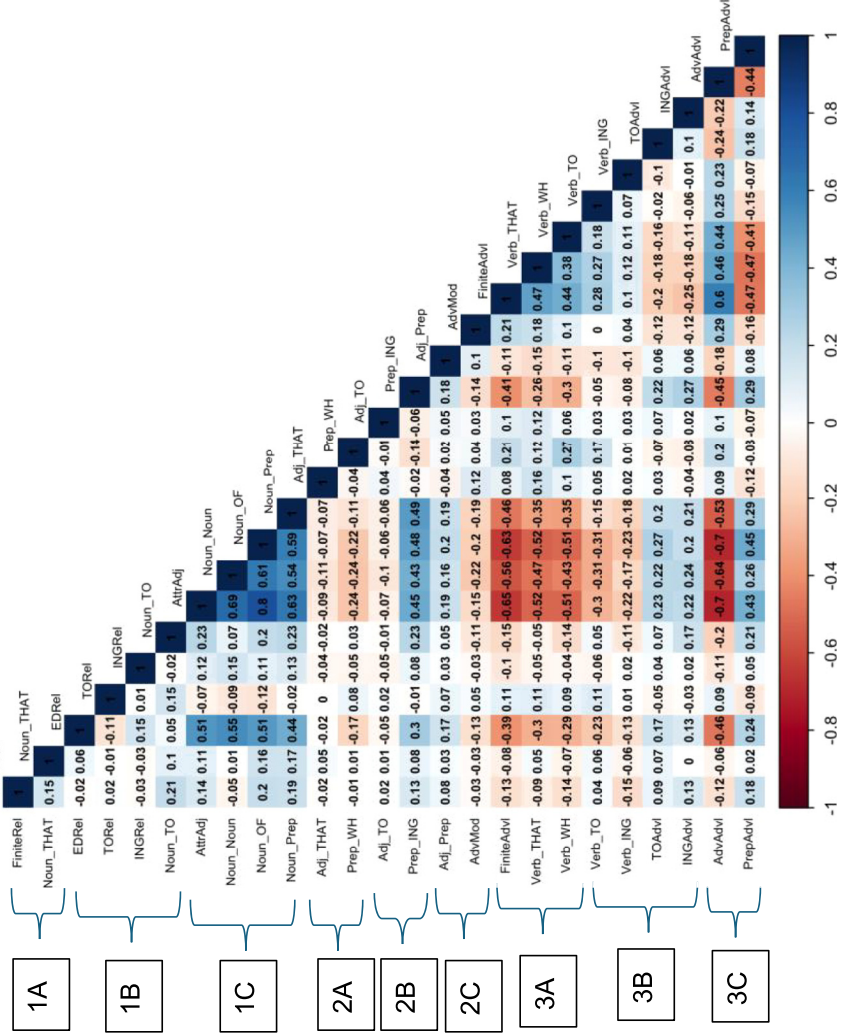

Figure 4: Heatmap of text-level correlations.

a few additional features that differ in either their structural or syntactic characteristics. We return to this finding in our discussion below.

Finally, the third major finding to note from Figure 4 is the large number of features that have no strong positive correlations with any other complexity feature. Thus, the following features have no correlations that are greater than +0.25 with any other complexity feature: FiniteRel, Noun + THAT, TORel, INGRel, Noun + TO, Adj + THAT, Prep + WH, Adj + TO, Adj + Prep, and Verb + ING.

In summary, three major patterns emerge[3] from a detailed inspection of Figure 4:

1. Confirmation of the two strong groupings of stereotypical phrasal complexity features (Cell 1C) and stereotypical clausal complexity features (Cell 3A), as well as the strong inverse relationship between those two cells.
2. Identification of three additional complexity features that have relatively strong co-occurrence relations with either the 1C features (EDRel and Prep + ING) or with the 3A features (AdvlAdv).
3. The absence of any strong text-level co-occurrence relations for the large majority of the other complexity features.

Apart from the new finding that three additional features (EDRel, Prep + ING, and AdvlAdv) pattern together with either the 1C grouping or the 3A grouping, the results of this exploratory analysis leave us with the same general conclusions that we had in the BLH papers, that: a) grammatical complexity in English is organized in terms of two strong groupings of complexity features: stereotypical phrasal features (1C) and stereotypical clausal features (3A); b) those two groupings occur in complementary distribution as a strong bi-polar dimension of variation (the 1C vs. 3A opposition); and c) "other features are distributed in their own peculiar ways, associated with their own idiosyncratic discourse functions" (Biber et al. 2024b: 92).

However, it turns out that this conclusion is premature. That is, the analyses to this point are all based on the assumption that functionally related features will co-occur regularly in the same texts. It might be the case, though, that features systematically pattern together in registers, even though they do not necessarily co-occur in the same texts. In the following section, we show that this is indeed the case, representing much stronger patterns of covariation than we could have anticipated based on the results presented in Figure 1. In fact, those patterns are so strong that they led us to reconsider the text-level results, finding (in Section 4 below)

3 To control for any possible bias from the corpus design, we ran an additional correlational analysis on a subcorpus that was balanced to contain the same number of texts for each register. Although the specific correlation coefficients changed somewhat, the patterns described here were exactly replicated in that analysis.

that the text-level patterns are entirely consistent with the register-level patterns, even though the strength of the correlations are much weaker.

3.2 Register-level correlations

Figure 5 has the same heatmap format as Figure 4, but it shows the extent to which each pair of complexity features systematically co-varies across registers (rather than across texts). The visual impact of Figure 5 is strikingly different from Figure 4: most of the cells in Figure 5 are intense blue or intense red (representing large positive or large negative correlations), in contrast to Figure 4, which had a predominance of white or lightly shaded cells (representing correlations closer to 0). Thus, the overall pattern shown in Figure 4 is that the majority of complexity features pattern together in highly systematic ways across registers, in contrast to their more idiosyncratic distributions across texts.

Similar to Figure 4, the strongest correlations shown on Figure 5 are among the 1C features and among the 3A features. All 1C features have correlations >0.85 with one another, while all 3A features have correlations >0.65 with one another. The inverse correlations between 3A features and 1C features are also quite remarkable, ranging from −0.64 to −0.91. Although our focus here is on the strength of these correlations, most of them are also statistically significant. (The critical value at $p < 0.05$ for a one-tailed test (i.e., hypothesizing a positive correlation) with $N = 8$ is $r = 0.549$.) Thus, the existence and strength of a higher-order bipolar dimension opposing stereotypical phrasal versus stereotypical clausal complexity features is strongly confirmed in Figure 5.

There are many additional cells in Figure 5 that show strong correlations. But surprisingly, similar to Figure 4, these other correlations do not indicate the existence of additional groupings beyond 1C and 3A. Rather, almost all of these other correlations occur between an additional feature and either the 1C grouping or the 3A grouping.

To better interpret the composition of these groupings, Figure 6 presents a heatmap based on the same bivariate correlations as in Figure 5, but reorganized to group together the features that have positive register-level correlations with 1C features (the upper-left triangle in the figure, shown in blue) versus the features that have positive register-level correlations with 3A features (the lower-right triangle in the figure, also shown in blue). The inverse correlations between these two groupings of features are shown in red in the lower-left rectangle of the figure. Figure 6 shows that eight additional features (EDRel, INGRel, Noun + TO, Prep + ING, Adj + Prep, TOAdvl, INGAdvl, PrepAdvl) have strong correlations with all 1C features (as well as generally strong correlations with one another), coupled with strong inverse

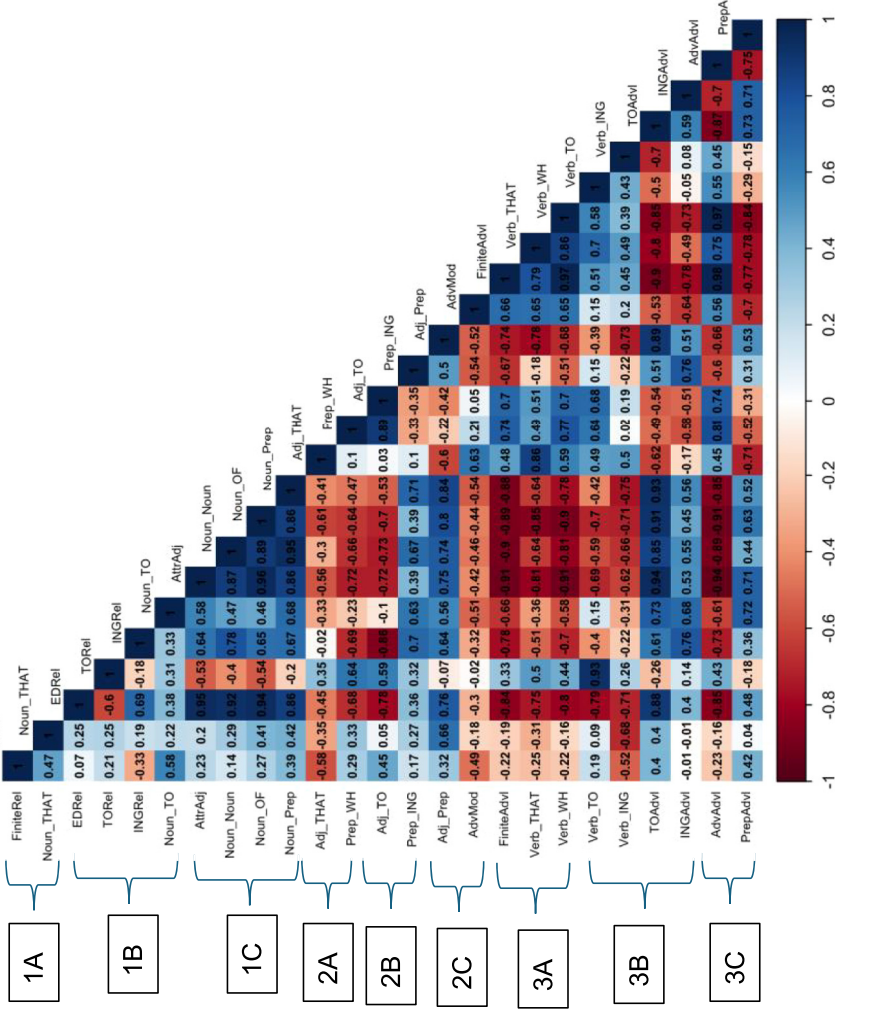

Figure 5: Heatmap of register-level correlations.

		Attr Adj	Noun _Noun	Noun _OF	Noun _Prep	EDRel	ING Rel	Noun _TO	Prep _ING	Adj _Prep	TO Advl	ING Advl	Prep Advl		Finite Rel	Noun _THAT		TO Rel	Adv Mod	Adj _TO	Prep _WH	Adj _THAT	Adv Advl	Verb _ING	Verb _TO	Finite Advl	Verb _THAT	Verb _WH
1C	AttrAdj	1																										
1C	Noun_Noun	0.87	1																									
1C	Noun_OF	0.96	0.89	1																								
1C	Noun_Prep	0.86	0.95	0.86	1																							
1B	EDRel	0.95	0.92	0.94	0.86	1																						
1B	INGRel	0.64	0.78	0.65	0.67	0.69	1																					
1B	Noun_TO	0.58	0.47	0.46	0.68	0.38	0.33	1																				
2B	Prep_ING	0.39	0.67	0.39	0.71	0.36	0.7	0.63	1																			
2C	Adj_Prep	0.75	0.74	0.8	0.84	0.76	0.64	0.56	0.5	1																		
3B	TOAdvl	0.94	0.85	0.91	0.93	0.88	0.61	0.73	0.51	0.89	1																	
3B	INGAdvl	0.53	0.55	0.45	0.56	0.4	0.76	0.68	0.76	0.51	0.59	1																
3C	PrepAdvl	0.71	0.44	0.63	0.52	0.48	0.36	0.72	0.31	0.53	0.73	0.71	1															
1A	FiniteRel	0.23	0.14	0.27	0.39	0.07	-0.33	0.58	0.17	0.32	0.4	-0.01	0.42		1													
1A	Noun_THAT	0.2	0.29	0.41	0.42	0.25	0.19	0.22	0.27	0.66	0.4	-0.01	0.04		0.47	1												
1B	TORel	-0.53	-0.4	-0.54	-0.2	-0.6	-0.18	0.31	0.32	-0.07	-0.26	0.14	-0.18		0.21	0.25		1										
2C	AdvMod	-0.42	-0.46	-0.44	-0.54	-0.3	-0.32	-0.51	-0.54	-0.52	-0.53	-0.64	-0.7		-0.49	-0.18		-0.02	1									
2B	Adj_TO	-0.72	-0.73	-0.7	-0.53	-0.78	-0.86	-0.1	-0.35	-0.42	-0.54	-0.51	-0.31		0.45	0.05		0.59	0.05	1								
2A	Prep_WH	-0.72	-0.66	-0.64	-0.47	-0.68	-0.69	-0.23	-0.33	-0.22	-0.49	-0.58	-0.52		0.29	0.33		0.64	0.21	0.89	1							
2A	Adj_THAT	-0.56	-0.3	-0.61	-0.41	-0.45	-0.02	-0.33	0.1	-0.6	-0.62	-0.17	-0.71		-0.58	-0.35		0.35	0.63	0.03	0.1	1						
3C	AdvAdvl	-0.94	-0.89	-0.91	-0.85	-0.85	-0.73	-0.61	-0.6	-0.66	-0.87	-0.7	-0.75		-0.23	-0.16		0.43	0.56	0.74	0.81	0.45	1					
3B	Verb_ING	-0.62	-0.66	-0.71	-0.75	-0.71	-0.22	-0.31	-0.22	-0.73	-0.7	0.08	-0.15		-0.52	-0.68		0.26	0.2	0.19	0.02	0.5	0.45	1				
3B	Verb_TO	-0.69	-0.59	-0.7	-0.42	-0.79	-0.4	0.15	0.15	-0.39	-0.5	-0.05	-0.29		0.19	0.09		0.93	0.15	0.68	0.64	0.49	0.55	0.43	1			
3A	FiniteAdvl	-0.91	-0.9	-0.89	-0.88	-0.84	-0.78	-0.66	-0.67	-0.74	-0.9	-0.78	-0.77		-0.22	-0.19		0.33	0.66	0.7	0.74	0.48	0.98	0.45	0.51	1		
3A	Verb_THAT	-0.81	-0.64	-0.85	-0.64	-0.75	-0.51	-0.36	-0.18	-0.78	-0.8	-0.49	-0.78		-0.25	-0.31		0.5	0.65	0.51	0.49	0.86	0.75	0.49	0.7	0.79	1	
3A	Verb_WH	-0.91	-0.81	-0.9	-0.78	-0.8	-0.7	-0.58	-0.51	-0.68	-0.85	-0.73	-0.84		-0.22	-0.16		0.44	0.65	0.7	0.77	0.59	0.97	0.39	0.58	0.97	0.86	1

r = .7 to 1.0
r = .3 to .69
r = .1 to .29
r = -.7 to -1.0
r = -.3 to -.69
r = -.1 to -.29
r is contrary to predicted pattern

Figure 6: Heatmap of register-level correlations, reorganized by literate versus oral complexity features.

correlations with all 3A features. An additional two features (FiniteRel and Noun + THAT) show the same pattern, but with weaker correlations. Conversely, Figure 6 further shows that eight additional features (Verb + TO, Verb + ING, AdvAdvl, Adj + THAT, Prep + WH, Adj + TO, AdvMod, TORel) have strong correlations with all 3A features (as well as generally strong correlations with one another), coupled with strong inverse correlations with all 1C features.

In summary, the register-level correlational analysis shows that:

a. the system of complexity features – at the register level – is organized almost entirely in terms of two major groupings of complexity features:
   1) a large group of complexity features that correlate with the 1C features, and
   2) a large group of complexity features that correlate with the 3A features.
b. and that those two groupings occur in complementary distribution with one another.

These two groupings cannot be interpreted as either a simple structural distinction between phrasal versus clausal features, or as a simple syntactic distinction between noun-modifiers versus clause modifiers/complements. Rather, both groupings include phrases as well as dependent clauses; both groupings include noun modifiers as well as clause-level constituents. Thus, the functional interpretation of these groupings is more complex than previously thought. We turn to that interpretation in the following section.

4 Interpretation and follow-up analyses

The results presented in Figure 6 show that complexity features at the register level are distributed as two major groupings of covarying features. We interpret these as the "literate complexity dimension" and the "oral complexity dimension." The features that comprise the literate dimension have positive register-level correlations with all 1C features, and in this sense, the 1C features can be regarded as the locus of the dimension. Similarly, the 3A features can be regarded as the locus of the oral dimension.

Structural type and syntactic function are very important considerations for the interpretation of these dimensions, even though they are not sufficient in themselves to completely account for the way in which the features covary. Table 3, which summarizes several major characteristics of each feature, shows that most features that are structural phrases belong to the literate dimension. The only exceptions are the phrasal features that involve adverbs – AdvMod and AdvAdvl – which belong to the oral dimension. In contrast, most finite dependent clause features belong to the oral dimension. The only exceptions are FiniteRel and Noun + THAT, which belong to

Table 3: Summary of the structural, syntactic, lexical, and functional characteristics for each complexity feature in the literate and oral dimensions.

Feature	Structural type (fin; nonfin; phrs)			Syntactic function (NMod, OtherMod, ClauseMod)			Lexically restricted			Communicative function (stance; InfoElab; other)		
	FinCls	NonFinCls	Phrs	ClsMod	OtherMod	NMod	Extreme	Moderate	Little	Stance	Other	Info
Literate features												
N + N			x			x			x			x
N + OF			x			x			x			x
Adj + N			x			x			x			x
N + Prep			x			x			x			x
EDRel		x				x			x			x
INGRel		x				x			x			x
N + TO		x				x	x			x		
Prep + ING		x		(x)	(x)	(x)			x			x
Adj + Prep			x	(x)	(x)			x				x
TOAdvl		x		x					x		x	
INGAdvl		x		x					x		x	
PrepAdvl			x	x					x			x
FiniteRel	x					x			x			x
N + THAT	x					x	x			x		
Oral features												
V + THAT	x			x			x			x		
V + WH	x			x				x			x	
FinAdvlCls	x			x			x			(x)	(x)	(x)
TORel		x				x	x					x
Adj + THAT	x			(x)	(x)		x			x		
Prep + WH	x			(x)	(x)	(x)			x		x	
Adj + TO		x		(x)	(x)		x			x		
AdvMod			x		x		x			x		
V + TO		x		x			x			(x)	(x)	
V + ING		x		x			x			(x)	(x)	
AdvAdvl			x	x			(x)	(x)		(x)	(x)	(x)

the literate dimension. Syntactic function is an equally important predictor of group membership: almost all features that function syntactically as noun modifiers belong to the literate dimension (including nonfinite clauses as well as finite clauses); the only exception is TORel. In contrast, most features that function as clause level constituents (adverbials or complements) belong to the oral dimension. The exceptions in this case are TOAdvl, INGAdvl, and PrepAdvl, which belong to the literate dimension.

Some of these features could be analyzed as serving multiple syntactic functions, which might help to explain their distributional patterns. For example, Adj + THAT and Adj + TO were classified as phrasal complements in the BLH studies (see Table 1; i.e., these are dependent clauses/phrases that complement an adjective head). However, if we consider the syntactic function of the higher-level construction (including the head adjective), these features could be regarded as clause-level complements, because the entire construction functions as the predicative of a copular verb.[4] For example:

I'm sure that they had two different reservations.
You're lucky to be alive.

If we reanalyze these features as having clause-level syntactic functions, their co-occurrence with other clause-level constituents is consistent with the overall linguistic interpretation of the oral dimension.

Altogether, the distributional patterning of most complexity features can be accounted for by these structural and syntactic principles, with only four features requiring additional explanation: TORels, Adj + TO, AdvAdvl, and AdvMod. Consideration of other functional characteristics can help with this additional explanation. Two of those characteristics, shown in Table 3, are the extent to which a complexity feature is lexically constrained, and the extent to which a complexity feature is associated with particular discourse functions (especially the expression

4 Similarly, features with a prepositional head (Prep + ING and Prep + WH) often function as clause-level constituents, although these same features can also function as a noun modifier:

Prep + ING and Prep + WH as clause-level constituents:

Flash can be used for building interactive multimedia applications.
You'll pay for what you do and say

Prep + ING and Prep + WH as noun modifier:

I am keen to see more opportunities for building stronger connections
Stevie Lis needs more recognition for what he's done

of stance) as opposed to more general informational functions. Overall Table 3 shows that these additional characteristics are strongly associated with the distinction between the oral and literate dimensions of complexity features: most features in the literate dimension are not lexically constrained, and they have general informational functions; in contrast, most features in the oral dimension are lexically constrained and serve stance discourse functions. At the same time, though, these characteristics help to account for the patterning of the four exceptional complexity features.

Some complexity features are lexically constrained, meaning that they occur with only a restricted set of controlling words. For example, there are only a few dozen nouns that can control a *that* noun complement clause (e.g., *fact, belief, possibility, assumption, claim*; see GSWE: 642–643), and a few dozen other nouns that can control a *to* noun complement clause (e.g., *attempt, effort, ability, opportunity, plan*; see GSWE: 646). In addition, other features are lexically constrained in actual use, even if they are less constrained in their potential. For example, there are over a hundred verbs that can control a verb complement clause, but only a few of those verbs are especially common (e.g., *think, say, know, see, find, believe, feel, suggest, show, guess*; see GSWE: 656–659).

Nearly all of the complexity features in the oral dimension are lexically constrained, including the four exceptional features identified above: TORels, Adj + TO, AdvAdvl, and AdvMod. For example, while there is a large set of possible nouns that can control a *to* relative clause, there is only a small set of nouns that frequently occur with this function, including *thing, person, time, place, stuff,* and *way* (see GSWE: 627–628). Similarly, there is a small set of adjectives that can control a *to* complement clause, and of those, only a few occur frequently (e.g., *(un)likely, difficult, easy, glad, hard, ready, (un)willing*; GSWE: 708–711). There is a much larger number of adverbs that can occur as adverbials, but in actual use, a relatively small set of those occur frequently. These include additive adverbs like *just, only, also, even*; time adverbs like *then, now, never, again, always, today*; place adverbs like *here* and *there*; and stance adverbs like *probably, maybe, perhaps, really,* and *actually* (see GSWE, pp. 788–793, 860–863). And finally, there is a relatively small set of degree adverbs that can modify another adverb or adjective, and an even smaller set of adverbs used frequently with this function (including amplifiers like *very, so, really, real, extremely, highly,* and hedges like *quite, pretty, relatively*; see GSWE: 560–563).

Table 3 also indicates the primary discourse function of each complexity feature. Much more work could be done on this consideration, but the major distinction here is between features that serve stance functions and features that serve general informational functions. Many of the complexity features in the oral dimension serve stance functions, while only two features in the literate dimension have this

primary function. This factor might be most important in accounting for the grouping of AdvMod with the oral features. That is, based on syntactic function (modifying adjective/adverb phrases) and structural type (a type of phrase), we might predict that AdvMods would group with the literate complexity features. But the combination of lexical restrictedness and stance functions seem to be major factors influencing the patterning of this feature with other oral features.

Of course, to fully interpret the nature of these two dimensions, we need to determine which registers they are associated with. That is, the correlation heatmaps show that complexity features covary with one another, but they do not tell us whether a feature is common or rare in a given register. For that goal, we need to analyze the rates of occurrence of each feature in each register.

Figure 7 presents those register patterns in the form of a heatmap. Unlike the correlational heatmaps presented in Figures 4–6, the cell values in Figure 7 are the Cohen's *d* scores for each feature in each register. Cohen's *d*, which has been employed to identify grammatical *key features* in a register (see Egbert and Biber 2023), shows the relative frequency of each linguistic feature in each register. Specifically, Cohen's *d* is the difference between a register mean and the overall corpus mean (of all registers taken together), measured in pooled standard deviation units. The advantage of using Cohen's *d* rather than absolute normalized frequencies is that

	Conversation	Classroom	Lectures	Fiction	Blogs	News	Textbooks	Research
AttrAdj	-1.20	-0.60	-0.13	-0.22	0.03	0.80	1.63	1.46
Noun_Noun	-0.92	-0.45	-0.48	-0.86	0.01	0.64	0.47	1.25
Noun_OF	-1.24	-0.49	0.33	-0.35	-0.03	0.52	1.29	1.62
Noun_Prep	-0.88	-0.14	-0.18	-0.94	0.33	0.67	0.63	0.98
EDRel	-0.68	-0.26	-0.28	-0.53	-0.14	0.14	0.89	0.90
INGRel	-0.10	-0.19	-0.18	-0.03	0.09	-0.03	0.00	0.24
Noun_TO	-0.39	-0.17	-0.05	-0.13	0.45	0.58	0.53	-0.01
Prep_ING	-0.87	-0.80	-0.84	-0.83	0.86	0.80	-0.84	0.90
Adj_Prep	-0.32	0.07	0.09	-0.22	0.22	0.00	0.35	0.41
TOAdvl	-0.36	-0.05	0.01	-0.16	0.16	0.21	0.51	0.40
INGAdvl	-0.49	-0.31	-0.39	0.28	0.26	0.28	0.11	0.36
PrepAdvl	-0.81	-0.10	0.30	1.02	0.61	0.97	1.19	0.40
FiniteRel	-0.42	0.45	0.90	-0.32	0.27	0.97	0.52	-0.27
Noun_THAT	-0.17	0.00	0.41	-0.42	0.21	-0.11	0.01	0.19
TORel	0.04	0.03	0.02	-0.11	0.30	0.02	-0.28	-0.35
AdvMod	0.34	-0.31	-0.12	-0.29	-0.25	-0.51	-0.02	-0.31
Adj_TO	0.03	0.16	0.11	0.00	0.03	0.04	-0.06	-0.21
Prep_WH	0.13	0.61	0.43	-0.35	0.22	-0.28	-0.33	-0.69
Adj_THAT	0.12	-0.16	-0.19	-0.11	-0.02	-0.08	-0.24	-0.09
AdvAdvl	0.86	0.76	0.11	-0.08	-0.17	-1.12	-1.01	-1.49
Verb_ING	0.20	-0.20	-0.36	0.67	0.03	-0.12	-0.47	-0.46
Verb_TO	0.22	-0.11	0.15	-0.11	0.55	0.11	-0.63	-0.83
FiniteAdvl	0.77	0.43	0.22	-0.09	-0.37	-0.88	-0.77	-1.31
Verb_THAT	0.72	-0.25	-0.39	-0.52	-0.16	-0.21	-1.00	-1.02
Verb_WH	0.56	0.37	-0.04	-0.36	-0.16	-0.66	-0.75	-1.01

Figure 7: Heatmap of Cohen's *d* values for the rate of occurrence of each complexity feature in each register.

feature rates can be directly compared for their importance across registers, regardless of how common the feature is in absolute terms. For example, Figure 7 shows a *d* value of 1.25 for N + N in research articles. That value shows that the mean of N + N in research articles (54.9 per 1,000 words) is 1.25 standard deviation units above the overall corpus mean of 26.6. *d* values do not reflect the absolute frequency of a feature; rather, they reflect the relative frequency of a feature in one register compared to the overall mean frequency in the entire corpus. EDRel has a *d* value of 0.9 in research articles, similar to the value of 1.25 for N + N. The actual rate of occurrence for EDRels in research articles is only 1.1 per 1,000 words – much less frequent than the rate of 54.9 per 1,000 words for N + N. However, EDRels have a large *d* score in research articles because that rate of occurrence is almost 1 standard deviation unit over the overall corpus mean of 0.4. Thus, both N + N and EDRels are strongly associated with research articles, even though the absolute frequencies of the two features are dramatically different.

Cohen's *d* values can also be negative, meaning that the rate of occurrence for a feature in a register is less than the overall corpus mean. So, for example, Figure 7 shows that N + N has a *d* value of -0.92 in conversation, meaning that the register mean of 10.9 per 1,000 words is 0.92 standard deviation units below the overall corpus mean of 26.6.

The features in the literate dimension are presented as the top 14 rows in Figure 7, and the features in the oral dimension are presented as the bottom 11 rows (following the same format as Figure 6). The overall patterns shown in the heatmap confirm the general expectations from previous register research:

- the literate complexity features tend to occur more frequently in written informational registers (especially research articles).
- the oral complexity features tend to occur more frequently in spoken registers (especially conversation).
- and the two dimensions of complexity features tend to be distributed in a complementary pattern, with the literate features being notably rare in spoken registers like conversation, and the oral features being notably rare in registers like written research articles.

Beyond those general patterns, there are several more specific trends that we can observe from Figure 7. Most literate complexity features are strongly and consistently associated with informational written registers. In addition, the literate features tend to occur with comparatively high frequencies even in less informational written registers like blogs. In contrast, the oral complexity features show more variation in their associations with spoken registers. For example, several oral features occur only slightly more frequently in conversation than the overall corpus mean (e.g., Adj + THAT, Prep + WH, Adj + TO, TORels). And Verb + THAT is quite

frequent in conversation but actually occurs less frequently than the corpus mean in spoken classroom teaching and lectures.

There are also specific findings that deserve more careful consideration. For example, fiction is a written register that patterns more similarly to spoken registers like conversation and lectures than it does to informational written registers. And features like FiniteRels and Noun + THAT have peculiar register patterns, different from the typical pattern seen for other literate features. For example, FiniteRels are especially frequent in newspaper writing and in spoken lectures, while they are actually somewhat rare in written research articles. Similarly, Noun + THAT are most frequent in spoken lectures, and moderately frequent in blogs and research articles. Specific patterns like these are consistent with the weaker correlations found for FiniteRels and Noun + THAT with other literate features (see Figure 6).

The discussion to this point has focused on the interpretation of the oral complexity dimension and the literate complexity dimension, as two independent parameters of linguistic variation. However, those dimensions should also be interpreted as constituting a single higher-order parameter of variation, because they occur in a strong complementary relationship to one another. That is, the two dimensions, working together, represent a fundamental choice in the way in which discourse is constructed. For example, Figure 7 shows that conversation employs the set of oral complexity features with high frequencies while at the same time it disprefers the use of all literate complexity features. And academic research articles show the opposite pattern, employing most literate complexity features with high frequencies while at the same time dispreferring the use of most oral complexity features. Thus, a register tends to rely on the oral set of complexity features, or on the literate set of complexity features, but not both. Some registers are more intermediate. For example, lectures generally rely on the oral set of complexity features, but not as frequently as conversation; lectures also disprefer the literate set of complexity features, but they use those features slightly more frequently than conversation.

Thus, this higher-order oral-versus-literate parameter of variation can be regarded as a continuum that represents a fundamental choice in discourse style. What we do not seem to find is a register that makes frequent use of both the oral and literate sets of features. However, it does seem possible for a register to make moderate use of both sets of complexity features. For example, Figure 7 shows that fiction employs most oral complexity features and most literate complexity with frequencies that are close to the overall corpus mean.

The major unresolved issue at this point is the question of why these patterns are observed so strongly at the register level (Figures 5 and 6), in contrast to our

conclusion in Section 3.1 above that no clear patterns could be observed at the text level (beyond the 1C and 3A goupings; Figure 4).[5] To further explore this question, we returned to the text-level correlations to analyze the extent to which they were consistent with the groupings of features in the register-level oral and literate dimensions. It turns out, if we focus on the direction of the correlation (positive or negative) rather than the magnitude of correlations, that the text-level findings shown in Figure 4 are actually consistent with the register-level findings shown in Figure 6. Figure 8 summarizes those patterns in the form of a heatmap, reorganizing the text-level correlations from Figure 4 to group together the literate features (the top 14 rows) versus the oral features (the bottom 11 rows). This heatmap shows only the correlations with the 1C and 3A features. The shading is designed primarily to capture positive correlations (shown in blue) versus negative correlations (shown in red), and so shading is used even in cases where the magnitude of the correlation is very small. At this level of analysis, almost all features in the literate dimension have positive text-level correlations with all 1C features, coupled with negative text-level correlations with all 3A features. And conversely, almost all features in the oral dimension have positive text-level correlations with all 3A features, coupled with negative text-level correlations with all 1C features. The fact that many of these correlations are weak in magnitude shows that many complexity features have their own peculiar distributions at the text level, a fact that requires further interpretation (see Section 6 below). However, the fact that the general patterning of positive versus negative correlations at the text level is nearly identical to that found at the register level provides strong confirmation for the importance of the oral versus literate dimensions of complexity features in language use.

5 Theoretical implications of the methodological approach adopted here

The register-level analyses in Sections 3.2 and 4 above represent a methodological departure from previous research carried out within the Text-Linguistic approach to

5 One source of low correlations is the simple fact that some of these grammatical features are rarely used, and thus many texts have no instances of those features. For example, *that* noun complement clauses never occur in 50 % of the texts in our corpus. Nonfinite *-ed* relative clauses never occur in 55 % of the texts, and *to* relative clauses never occur in 70 % of the texts. Because these features occur with a rate of 0.0 per 1,000 words in those texts, it is not surprising that the features would have at best weak correlations with other features at the text level. However, when features are analyzed at the register level, even these rare features have some rate of occurrence in every register, making it possible for them to have more meaningful correlations with other features.

		Attr Adj	Noun _Noun	Noun _OF	Noun _Prep	EDRel	ING Rel	Noun _TO	Prep _ING	Adj _Prep	TO Advl	ING Advl	Prep Advl	Finite Rel	Noun _THAT	TO Rel	Adv Mod	Adj _TO	Prep _WH	Adj _THAT	Adv Advl	Verb _ING	Verb _TO	Finite Advl	Verb _THAT
1C	AttrAdj	1																							
1C	Noun_Noun	0.69	1																						
1C	Noun_OF	0.8	0.61	1																					
1C	Noun_Prep	0.63	0.54	0.59	1																				
1B	EDRel	0.51	0.55	0.51	0.44																				
1B	INGRel	0.12	0.15	0.11	0.13																				
1B	Noun_TO	0.23	0.07	0.2	0.23																				
2B	Prep_ING	0.45	0.43	0.48	0.49																				
2C	Adj_Prep	0.19	0.16	0.2	0.19																				
3B	TOAdvl	0.23	0.22	0.27	0.2																				
3B	INGAdvl	0.22	0.24	0.2	0.21																				
3C	PrepAdvl	0.43	0.26	0.45	0.29																				
1A	FiniteRel	0.14	-0.05	0.2	0.19																				
1A	Noun_THAT	0.11	0.01	0.16	0.17																				
1B	TORel	-0.07	-0.09	-0.12	-0.02																				
2C	AdvMod	-0.15	-0.22	-0.2	-0.19																				
2B	Adj_TO	-0.07	-0.1	-0.06	-0.06																				
2A	Prep_WH	-0.24	-0.24	-0.22	-0.11																				
2A	Adj_THAT	-0.09	-0.11	-0.07	-0.07																				
3C	AdvAdvl	-0.7	-0.64	-0.7	-0.53																				
3B	Verb_ING	-0.22	-0.17	-0.23	-0.18																				
3B	Verb_TO	-0.3	-0.31	-0.31	-0.15																				
3A	FiniteAdvl	-0.65	-0.56	-0.63	-0.46	-0.39	-0.1	-0.15	-0.41	-0.11	-0.2	-0.25	-0.47	-0.13	-0.08	0.11	0.21	0.1	0.21	0.08	0.6	0.1	0.28	1	
3A	Verb_THAT	-0.52	-0.47	-0.52	-0.35	-0.3	-0.05	-0.05	-0.26	-0.15	-0.18	-0.18	-0.47	-0.09	0.05	0.11	0.18	0.12	0.12	0.16	0.46	0.12	0.27	0.47	1
3A	Verb_WH	-0.51	-0.43	-0.51	-0.35	-0.29	-0.04	-0.14	-0.3	-0.11	-0.16	-0.11	-0.41	-0.14	-0.07	0.09	0.1	0.06	0.27	0.1	0.44	0.11	0.18	0.44	0.38

r = .40 to .90
r = .10 to .39
r = .01 to .09

r = -.40 to -.90
r = -.10 to -.39
r = -.01 to -.09

r is contrary to predicted pattern

Figure 8: Heatmap of text-level correlations with the 1C and 3A features, reorganized by literate versus oral complexity features.

register variation (Biber 2019). In that approach, analyses are based on the situational and linguistic characteristics of each text. Theoretically, this approach is motivated by the claim that language use occurs as texts. Texts are usually instances of a higher-level register, but speakers and writers produce language as texts, not registers. And quantitatively, the Text-Linguistic approach is motivated by the fact that it enables analyses of variation within a register, in addition to analyses of the typical characteristics of a register (see discussion in Biber 2019; Biber and Egbert 2023; Egbert et al. 2025).

All previous Multi-Dimensional (MD) studies of register variation have been based on a Text-Linguistic analysis (see the survey in Goulart and Wood 2021). In that body of research, *dimensions* are conceptualized as "different sets of co-occurring linguistic features [that reflect] different functional underpinnings (e.g., interactiveness, planning, informational focus)" (Biber 2019: 50). For our purposes here, the crucially important aspect of this definition relates to the definition of "co-occurrence" as sets of features that occur together in the same texts and covary in systematic ways across texts. That is, the MD research program has been based on the relatively strong assumption that sets of functionally related features (constituting "dimensions") can be identified empirically because they co-occur regularly in the same texts and, therefore, covary in systematic ways across a corpus of texts.

Research based on this assumption has proven to be highly productive. The dimensions identified in MD studies usually account for a large proportion of the linguistic variation in the target corpus. At the same time, though, there has always been considerable linguistic variation that is not accounted for by a standard MD analysis. For example, the shared variance underlying the 5 major dimensions extracted in the Biber (1988) study of spoken and written registers account for only 46.5 % of all variance among that set of linguistic features. Similarly, the shared variance underlying the 4 dimensions extracted in the Biber (2006) study of university registers account for only 46.9 % of all variance among that set of linguistic features. Thus, although these sets of co-occurring linguistic features (the "dimensions") have proven to be strong predictors of register variation, there has also always been a large pool of linguistic variation that is not accounted for in MD analyses.

A similar pattern was found in the BLH text-linguistic studies of complexity features: although the analyses (based on text-level rates of occurrence) accounted for a large pool of shared linguistic covariation, there was also considerable variation that was not accounted for. This finding led to the present study, employing the alternative register-level analysis to determine whether complexity features covary systematically across registers, even if the same features are not co-occurring regularly in texts. Future research is required to reconcile the findings from these

two approaches and to better develop the methods for register-level analyses. Specifically, we plan to explore additional functional considerations that might help to explain the patterns of variation for complexity features across texts, statistical methods (such as cluster analysis) that could help to identify groupings of linguistic features at the register level, and the applicability of register-level analyses in other cases where previous analyses have accounted for only a relatively small proportion of the total linguistic variation.

6 Summary and conclusion

So, what overall generalizations can be made from these results, and how do they relate to the findings and conclusions of previous studies? The results of the earlier BLH studies suggested that grammatical complexity might be organized as two basic dimensions of language use associated with the 1C and 3A cells, combined with a large number of additional complexity features that served their own specialized discourse functions. For example, in Biber et al. (2024a), we concluded that

> the results of the present study show that some groupings of complexity features pattern together as underlying dimensions of grammatical text complexity. Other complexity features have their own peculiar distributions, apparently motivated by specialized discourse functions. (BLH 2024a: 371)

In part, these conclusions are confirmed in the present study. Although the 1C and 3A groupings are strongly confirmed, the text-level analysis in Section 3.1 shows that many complexity features have only weak correlations with any other complexity feature. This finding reflects the fact that these other features have their own peculiar distributions across texts and, therefore, apparently serve their own distinct discourse functions. A complete discourse analysis of all structural/syntactic complexity features in the grammatical system of English is required to fully investigate the range of those specific discourse functions.

In contrast, though, the register-level results in Section 3.2 strongly indicate that all structural/syntactic complexity features pattern together with either the 1C or the 3A groupings. These larger groupings of features are interpreted here as the "literate complexity dimension" and the "oral complexity dimension." In addition, more detailed analysis (in Section 4) shows that these same two dimensions exist at the text level, even though the interfeature correlations are much weaker at that level.

Thus, we are left with two complementary perspectives that are both supported by empirical evidence: On the one hand, many complexity features are only weakly correlated with other complexity features at the text level, indicating that they have their own peculiar distributions across texts, and that they serve their own distinct discourse functions. At the same time, though, there is strong evidence that all complexity features in the grammatical system of English pattern together at the register level with either the oral complexity dimension or the literate complexity dimension. Specific features are more or less strongly associated with that underlying system, but the entire grammatical system of structural/syntactic complexity features is organized as two underlying dimensions of variation, referred to here as the "literate" complexity dimension and the "oral" complexity dimension.

In addition, the results here show that these two major complexity dimensions actually function together as opposite poles of a single higher-order parameter of variation. This conclusion is consistent with the results of the earlier BLH studies, which found that:

> the stereotypical clausal ("oral") complexity dimension [i.e., 3A features] and the stereotypical phrasal ("literate") complexity dimension [i.e., 1C features] have a systematic but complementary relation to each other: when a text frequently employs oral complexity features, that same text (or register) tends to rarely use literate complexity features, and vice versa. (BLH 2024a: 365).

Biber et al. (2024a) further tested the statistical strength of this complementary relationship, finding that "this specialized model exhibits perfect fit, strongly confirming the hypothesis that [the 1C and the 3A cells] represent a fundamental opposition between the typical linguistic styles of spoken discourse versus written discourse" (p. 366). The present study provides even stronger support for that relationship: at the register level, all complexity features pattern together as part of either the literate or the oral dimension (i.e., not only the 1C vs. the 3A features), and the complementary distribution of those dimensions extends to each of these features. That is, all literate complexity features have robust negative correlations with all oral complexity features, and vice versa.

Thus, the system of complexity features is organized in a more complex manner than we might have initially anticipated. On the one hand, grammatical complexity is organized as two separate underlying dimensions: the group of literate complexity features must be analyzed separately from the group of oral complexity features. At the same time, the cumulative research findings of previous research, coupled with the specific findings of the present study, clearly show

that those two dimensions function together as opposite poles of a single higher-order parameter of variation.

In practice, it is necessary to separately analyze the grammatical features associated with the oral dimension versus the literate dimension, and in this sense, complexity must be regarded as a two-dimensional construct. At the same time, though, we need to recognize the way in which those two complexity dimensions function as complementary components of an underlying higher-order parameter of variation. The oral and literate dimensions represent two fundamentally different ways of constructing discourse; the two occur as a complementary choice. Thus, these are not merely two independent dimensions; rather, they are opposing poles of a single higher-order parameter of variation.

It is important to emphasize the methodological implications of these findings for applied studies of grammatical complexity. In particular, we want to emphasize that the findings here provide absolutely no support for the use of omnibus measures (which combine occurrences of oral and literate complexity features). Rather, the use of oral complexity features must be analyzed separately from the use of literate complexity features: these are two fundamentally different dimensions of complexity. Those dimensions constitute a single higher-order parameter because they represent two complementary ways of constructing discourse. Thus, it would make no sense to add together occurrences of features from the two basic dimensions. Rather, texts must be analyzed for the extent to which they use oral features as opposed to literate features, and vice versa.

There is still much that we do not understand about the distribution and function of grammatical complexity features. For example, future research is required to investigate a wider range of spoken and written registers, as well as the way in which the use of these complexity features is acquired by language learners. We also need additional research to more fully understand the interplay between the specific functions and distributions of individual complexity features across texts versus the shared functions and distributions of the literate/oral sets of features. We are very interested in the cross-linguistic generalizability of these findings: Is the contrast between oral versus literate complexity features a cross-linguistic universal of register variation, or are these patterns restricted to English (compare Biber 2014). And finally, we need additional research to fully understand the underlying motivations for these patterns. That is, while we can provide strong empirical evidence for the existence of the literate complexity dimension, the oral complexity dimension, and the literate-versus-oral parameter of variation, we have a less complete understanding of why these particular features should pattern together in this way. These are the questions that we hope to explore further in future research.

Appendix: Phrasal/clausal complexity features in English, by structural type and syntactic function

Complexity feature	Structural type	Syntactic function	Examples
Causative, conditional, concessive clause	Finite Dependent Clause	Clause constituent: Adverbial	*She won't narc on me, because she prides herself on being a gangster. Well, if I stay here, I'll have to leave early in the morning.*
Verb + *that* complement clause	Finite Dependent Clause	Clause constituent: Verb complement	*I would hope that we can have more control over them*. (with ZERO complementizer): *yeah, I think I probably could*.
Verb + *wh* complement clause	Finite Dependent Clause	Clause constituent: Verb complement	*I don't know how they do it*.
Noun + Finite relative clause (*that* or *WH*)	Finite Dependent Clause	Noun phrase constituent: NP modifier	*…the experimental error that could result from using cloze tests*
Noun + *that* complement clause	Finite Dependent Clause	Noun phrase constituent: NP complement	*The fact that no tracer particles were found indicates that these areas are not a pathway…*
Adjective + *that* complement clause	Finite Dependent Clause	Other phrase constituent: Adjective complement	*We're happy that the hunger strike has ended*.
Extraposed adjective + *that* complement clause	Finite Dependent Clause	Other phrase constituent: Adjective complement	*It is evident that the virus formation is related to the cytoplasmic inclusions*.
Preposition + *wh* complement clause	Finite Dependent Clause	Other phrase constituent: Prepositional complement	*I'll offer a suggestion for what we should do*.
to-clause as "purpose" adverbial	Nonfinite Dependent Clause	Clause constituent: Adverbial	*To verify this hypothesis, sections of fixed cells were examined.*
ing-clause as adverbial	Nonfinite Dependent Clause	Clause constituent: Adverbial	*Considering mammals' level of physical development, the diversity of this species is astounding.*

(continued)

Complexity feature	Structural type	Syntactic function	Examples
ed-clause as adverbial	Nonfinite Dependent Clause	Clause constituent: Adverbial	*Based on estimates of the number of unidentified species, other studies put the sum total in the millions.*
Verb + *to* complement clause	Nonfinite Dependent Clause	Clause constituent: Verb complement	*I really want to fix this room up.*
Verb + *ing* complement clause	Nonfinite Dependent Clause	Clause constituent: Verb complement	*I like watching the traffic go by.*
Noun + *-ed* (passive) relative clause	Nonfinite Dependent Clause	Noun phrase constituent: NP modifier	*This is a phrase used in the recruitment industry.*
Noun + *-ing* relative clause	Nonfinite Dependent Clause	Noun phrase constituent: NP modifier	*Elevated levels are treated with a diet consisting of low cholesterol foods.*
Noun + *to* relative clause	Nonfinite Dependent Clause	Noun phrase constituent: NP modifier	*You're the best person to ask.*
Noun + *to* complement clause	Nonfinite Dependent Clause	Noun phrase constituent: NP complement	*The project is part of a massive plan to complete the section of road…*
Adjective + *to* complement clause	Nonfinite Dependent Clause	Other phrase constituent: Adjective complement	*I was happy to do it.*
Extraposed adjective + *to* complement clause	Nonfinite Dependent Clause	Other phrase constituent: Adjective complement	*It was important to obtain customer feedback.*
Preposition + *ing* complement clause	Nonfinite Dependent Clause	Other phrase constituent: Prepositional complement	*The formula for calculating the effective resistance is ….*
Adverb phrase as adverbial	Dependent phrase (nonclausal)	Clause constituent: Adverbial	*I raved about it afterwards.*
Prepositional phrase as adverbial	Dependent phrase (nonclausal)	Clause constituent: Adverbial	*Alright, we'll talk to you in the morning.*

(continued)

Complexity feature	Structural type	Syntactic function	Examples
Attributive adjectives as noun premodifier	Dependent phrase (nonclausal)	Noun phrase constituent: NP modifier	*emotional injury*, *conventional practices*
Nouns as noun premodifier	Dependent phrase (nonclausal)	Noun phrase constituent: NP modifier	*aviation security committee*, *fighter pilot training*
Of genitive phrases as noun postmodifier	Dependent phrase (nonclausal)	Noun phrase constituent: NP modifier	*McKenna wrote about the origins of human language*.
Other prepositional phrases as noun postmodifier	Dependent phrase (nonclausal)	Noun phrase constituent: NP modifier	*Overall scores were computed by averaging the scores for male and female students*.
Appositive noun phrases as noun postmodifier	Dependent phrase (nonclausal)	Noun phrase constituent: NP modifier	*James Klein, president of the American Benefits Council*
Prepositional phrases as adjective complement	Dependent phrase (nonclausal)	Other phrase constituent: Adjective complement	*I'd be happy with just one*.
Adverb phrase as adjective/adverb modifier	Dependent phrase (nonclausal)	Other phrase constituent: Adjective/ adverb modifier	*That cat was surprisingly fast.* *We will see those impacts fairly quickly.*

References

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

Biber, Douglas. 2006. *University language: A corpus-based study of spoken and written registers*. Amsterdam: John Benjamins.

Biber, Douglas. 2014. Using multi-dimensional analysis to explore cross-linguistic universals of register variation. *Languages in Contrast* 14(1). 7–34.

Biber, Douglas. 2019. Text-linguistic approaches to register variation. *Register Studies* 1. 42–75.

Biber, Douglas & Jesse Egbert. 2023. What is a register? Accounting for linguistic and situational variation within – and outside of – textual varieties. *Register Studies* 5. 1–22.

Biber, Douglas, Bethany Gray, Shelley Staples & Jesse Egbert. 2020. Investigating grammatical complexity in L2 English writing research: Linguistic description versus predictive measurement. *Journal of English for Academic Purposes* 46. 1–14.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 2021. *The grammar of spoken and written English*. Amsterdam: John Benjamins [Previously published as *The Longman grammar of spoken and written English*, 1999].

Biber, Douglas, Bethany Gray, Shelley Staples & Jesse Egbert. 2022. *The register-functional approach to grammatical complexity: Theoretical foundation, descriptive research findings, applications*. London: Routledge.

Biber, Douglas, Larsson Tove & Gregory R. Hancock. 2024a. The linguistic organization of grammatical text complexity: Comparing the empirical adequacy of theory-based models. *Corpus Linguistics and Linguistic Theory* 20(2). 347–373.

Biber, Douglas, Larsson Tove & Gregory R. Hancock. 2024b. Dimensions of text complexity in the spoken and written modes: A comparison of theory-based models. *Journal of English Linguistics* 52(1). 65–94.

Biber, Douglas, Larsson Tove, Gregory R. Hancock, Randi Reppen, Shelley Staples & Bethany Gray. 2025a. Comparing theory-based models of grammatical complexity in student writing. *International Journal of Learner Corpus Research* 11. 145–177.

Biber, Douglas, Bethany Gray, Tove Larsson & Shelley Staples. 2025b. Grammatical analysis is required to describe grammatical (and "syntactic") complexity. *Language Learning*. https://doi.org/10.1111/lang.12683.

Bulté, Bram & Alex Housen. 2018. Syntactic complexity in L2 writing: Individual pathways and emerging group trends. *International Journal of Applied Linguistics* 28. 147–164.

Bulté, Bram, Alex Housen & Gabriele Palotti. 2025. Complexity and difficulty in second language acquisition: A theoretical and methodological overview. *Language Learning*. https://doi.org/10.1111/lang.12669.

Carter, Ron & Michael McCarthy. 2006. *Cambridge grammar of English*. Cambridge: Cambridge University Press.

Casal, J. Elliott & Joseph J. Lee. 2019. Syntactic complexity and writing quality in assessed first-year L2 writing. *Journal of Second Language Writing* 44. 51–62.

Egbert, Jesse & Douglas Biber. 2023. Key feature analysis – A simple, yet powerful method for comparing text varieties. *Corpora* 18(1). 121–133.

Egbert, Jesse, Douglas Biber, Marianna Gracheva & Daniel Keller. 2025. Register and the dual nature of functional correspondence: Accounting for text-linguistic variation between registers, within registers, and without registers. *Corpus Linguistics and Linguistic Theory*. https://www.degruyterbrill.com/document/doi/10.1515/cllt-2024-0011/html.

Goulart, Larissa & Margaret Wood. 2021. Methodological synthesis of research using multi-dimensional analysis. *Journal of Research Design and Statistics in Linguistics and Communication Science* 6(2). 107–137.

Huddleston, Rodney. 1984. *Introduction to the grammar of English*. Cambridge: Cambridge University Press.

Huddleston, R. & G. K. Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.

Kushik, Ghulam A. & Ari Huhta. 2020. Investigating syntactic complexity in EFL learners' writing across common European framework of reference levels A1, A2, and B1. *Applied Linguistics* 41. 506–532.

Lu, Xiaofei. 2017. Automated measurement of syntactic complexity in corpus-based L2 writing research and implications for writing assessment. *Language Testing* 34. 493–511.

Purpura, James E. 2004. *Assessing grammar*. Cambridge: Cambridge University Press.

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

Szmrecsanyi, Benedikt. 2015. Recontextualizing language complexity. In Jocelyne Daems, Eline Zenner, Kris Heylen, Dirk Speelman & Hubert Cuyckens (eds.), *Change of paradigms: New paradoxes: Recontextualizing language and linguistics. Applications of Cognitive Linguistics 31*, 347–360. Berlin: De Gruyter Mouton.

Willis, David. 2003. *Rules, patterns and words: Grammar and lexis in English language teaching*. Cambridge: Cambridge University Press.