OPEN PEER COMMENTARY

Grammatical Analysis Is Required to Describe Grammatical (and "Syntactic") Complexity: A Commentary on "Complexity and Difficulty in Second Language Acquisition: A Theoretical and Methodological Overview"

Douglas Biber,[a] Bethany Gray,[b] Tove Larsson [iD],[a]
and Shelley Staples[c]

[a]Northern Arizona University [b]Iowa State University [c]University of Arizona

Our response is focused on the analysis of grammatical complexity[1] in learner language. We fully support the distinction between complexity and difficulty advocated by Bulté et al., and we agree with their characterization of "complexity" (as opposed to "difficulty") as referring "to the structural characteristics of linguistic items/structures and texts." We are surprised, though, by Bulté et al.'s recommendation that this construct of grammatical complexity can be adequately captured by omnibus measures that disregard and confound the different influences of multiple structural and syntactic considerations. That is, even though Bulté et al. recommend using multiple omnibus measures that are intended to capture different "subdimensions" of complexity, all of those measures are extremely general, disregarding analysis of particular grammatical structures and completely disregarding analysis of syntactic function. Thus, the methods recommended by Bulté et al. (and those generally practiced by second language acquisition [SLA] researchers) are based on the implicit assumption that grammatical complexity can be

---

Correspondence concerning this article should be addressed to Douglas.Biber@NAU.EDU
The handling editor for this article was Scott Jarvis.

described without actually carrying out a careful grammatical/syntactic analysis. As opposed to that approach, we argue that any adequate description of grammatical complexity must be based on a principled linguistic analysis of grammatical structures and syntactic functions.

To illustrate how anomalous the "omnibus" approach is, imagine a team of biologists who want to describe and compare the complexity of forests. These researchers are aware of the incredible diversity in the composition of forests. For example, one forest is composed of deciduous and coniferous trees from many different species, at different stages of maturity, growing with different extents of density, with undergrowth representing many different plant species; another forest is composed entirely of pine trees all at a single stage of maturity with no undergrowth. However, the researchers decide to disregard what they know about the biology of trees and plants, and instead simply operationalize "forest complexity" as the average height of trees in a forest and the mean number of branches per tree—focusing on only two general characteristics of forests while disregarding numerous characteristics that make forests fundamentally different from one another. It doesn't seem credible that a specialist from biology would advocate such a reductionist approach. Surprisingly, though, this example is similar to the widespread practice of SLA researchers (including Bulté et al.) who analyze grammatical complexity through the use of omnibus measures like mean number of words per phrase or mean number of clauses per T-unit.

We agree with Bulté et al.'s definition of grammatical complexity as "the quantity and variety of constituents and relationships between constituents." This definition comprises two main considerations: the nature of the constituents (the different grammatical *structures*) and the relationships between constituents (what we refer to as their *syntactic functions*). However, the omnibus-measure approach that Bulté et al. recommend is completely at odds with their definition: The reliance on omnibus measures fails to distinguish systematically among different grammatical structures, and it mostly disregards the analysis of different syntactic functions.[2]

Our argument is based on the fact that grammatical complexity in English[3] is itself an incredibly complex linguistic 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., Quirk et al., 1985; Biber et al., 1999/2021; Huddleston & Pullum, 2002).

It is also important to note that there is no precedent in previous linguistic research or theory for the practice of disregarding or collapsing consideration of grammatical/syntactic distinctions. In fact, ever since the 1960s, the sub-disciplines of variationist linguistics (e.g., Labov, 1969; Szmrecsanyi, 2017) and functional linguistics (e.g., Nichols, 1984) have been based on the fundamental premise that all linguistic variation is meaningful and therefore must be accounted for.

So, it seems intuitively obvious to us that any analysis of grammatical complexity in English learner language would necessarily be based on analysis of the grammatical structures and syntactic functions that comprise the grammatical system of English. Surprisingly, though, this has not been standard practice. Rather, SLA researchers commonly rely on omnibus measures that confound analysis of different grammatical structures and completely disregard analysis of syntactic function—an approach that Bulté et al. continue to advocate for.[4]

To illustrate the problems with omnibus measures, consider the analysis of two sentences discussed in Biber et al. (2022, pp. 435–436), one from a spoken interview (7) and the other from a medical news article (8).

7. The thing I believed would never happen was that Christa was told that she would probably never be able to have a child.
8. There is a need for further high quality research into the association between the experience of stress across a variety of contexts and miscarriage risk.

These two sentences have nearly identical values for the omnibus measure of *T-unit length*:[5] 23 words in Sentence 1, and 25 words in Sentence 2. But the two sentences illustrate dramatically different structural and syntactic characteristics. For example, Sentence 7 shows extensive embedding with five different dependent clauses, including several structural types serving different syntactic roles: a finite relative clause as noun modifier (*the thing **I believed** …*), three finite complement clauses controlled by verbs (*believed **[the thing] would never happen***; *was **that Christa was told***; *was told **that she would never be able …***), and a nonfinite complement clause controlled by an adjective (*be able **to have a child***).

In contrast, Sentence 8 has no embedded dependent clauses at all. Instead, this T-unit is long because there are several embedded prepositional phrases (marked in (9) below with square brackets) that modify head nouns (shown in bold):

9. There is a **need** [A for further high quality **research** [B into the **association** [C between the experience of **stress** [D across a variety of contexts D] and miscarriage risk C] B] A]

These examples illustrate the basic characteristic of all omnibus measures: They collapse the influence of multiple structural types and syntactic functions into a single number, and as a result, fundamentally different types of grammatical complexity are incorrectly characterized as being the same.

Bulté et al. continue to recommend the use of omnibus measures (e.g., "clauses per T-unit"), including new measures that (to our knowledge) have not been previously used in SLA research (e.g., "words per phrase" and "phrases per clause"). In addition, in a major departure from previous SLA research, Bulté et al. propose the use of a measure focused on syntactic function ("MATTR of dependency relations or syntactic structures").[6] For our purposes here, though, the important characteristic of all of these measures is that they are omnibus variables that confound the influence of multiple types of

complexity structures and disregard the influence of different syntactic functions. For example, the length of a phrase gives no consideration to the structural type of that phrase (e.g., noun phrase, adjective phrase, prepositional phrase), and no consideration to the types of structure that cause the phrase to become longer (e.g., embedded adjectives versus embedded prepositional phrases versus embedded nonfinite clauses versus embedded finite dependent clauses). In addition, most of these measures completely disregard the role of syntactic function. The MATTR measure is exceptional in that it focuses on syntactic function, but it is still an omnibus measure that simply captures the number of different syntactic functions; that is, this measure fails to capture the use of particular syntactic functions (and it disregards structural distinctions). In summary, while Bulté et al. pay attention to the representation of different "subdimensions of complexity," the specific measures that they propose continue to suffer from the problems of other omnibus measures used in SLA research.

In previous publications, we have presented empirical corpus-based evidence showing that the structural and syntactic distinctions found in the complexity system of English actually matter: They are extremely important in descriptions of language use for capturing the differing kinds of complexity in different registers (see, e.g., Biber et al., 2022; Biber, Larsson, & Hancock 2024a, 2024b; Biber, Larsson, Hancock, Reppen, et al., 2024). For example, finite dependent clauses, functioning syntactically as clause-level constituents (e.g., *I don't know **how they do it***), are much more common in conversational spoken registers than in written academic registers. In contrast, phrases functioning syntactically as noun phrase modifiers (e.g., *scores **for male and female students***) are much more common in written academic registers. A study based on omnibus measures—which would confound those structural distinctions and completely disregard those syntactic distinctions—would fail to capture those fundamentally important patterns of use. For the same reasons, complexity studies of English learner language based on omnibus measures fail to adequately reflect how learners are actually using the resources found in the complexity system of English.

However, we want to emphasize in closing that our argument here is based on the grammatical system of complexity features in English. That is, although our previous corpus-based research provides strong empirical support for the importance of these structural/syntactic distinctions, we are arguing here—on purely linguistic grounds—that any adequate description of grammatical complexity in English learner language must take account of the full range of different structures *and* syntactic functions that comprise

the grammatical system of English. No researcher would ignore their years of training in biology if they were describing the complexity of a forest. We similarly believe that no researcher should ignore their years of training in linguistics when describing the grammatical complexity of a text. We recognize that such analyses will require more work than automatically computing a value for an omnibus measure. But that is not sufficient reason to eschew them. Rather, we encourage collaboration between SLA researchers and grammarians to develop methods that are linguistically adequate while at the same time feasible for the practicing SLA researcher.

<div align="right">Final revised version accepted 24 August 2024</div>

Notes

1 It is surprising to us that researchers in second language acquisition (SLA) strongly prefer the label "syntactic complexity" as opposed to "grammatical complexity"—even though almost all SLA studies of complexity completely disregard analysis of syntactic function, focusing instead solely on structural characteristics. We explain this distinction further in our response below.

2 Bulté et al. do recommend one measure focused on the number of different syntactic functions in a text ("MATTR of dependency relations or syntactic structures"). We see that recommendation as an advance over previous SLA research, which has completely disregarded any analysis of syntactic function. However, this is still an omnibus measure, which indicates only the number of different syntactic functions employed by a student, regardless of what those functions are, and regardless of the extent to which particular structures are used for particular syntactic functions.

3 Similar arguments could be made for the analysis of grammatical complexity in any other language.

4 From a historical perspective, it is interesting to note that Hunt (1965)—the creator of T-unit based omnibus measures—devoted seven chapters to the discussion of particular grammatical structures serving particular syntactic functions, in contrast to only two chapters devoted to the discussion of omnibus measures based on general T-unit characteristics.

5 T-unit length is probably the most widely utilized omnibus measure in SLA complexity research.

6 Although Bulté et al. do not discuss the theoretical motivation for these new omnibus measures, we are encouraged by the fact that they seem to recognize the importance of distinguishing between phrasal versus clausal complexity, and the importance of distinguishing between structural complexity versus syntactic-function complexity. However, these new variables are still omnibus measures, suffering from the analytical shortcomings described above. In addition, Bulté et al. fail to provide operational definitions for these new measures, making it impossible to evaluate them in detail or to employ them in practice. For example, to

determine values for "words per phrase" and "phrases per clause," we need operational definitions of "phrase" and "clause," and we need linguistically sound methods for dealing with phrases that are embedded in higher level phrases. So, for example, does the length of a phrase include the lengths of all embedded phrases? And then should we subsequently compute the lengths of each of those embedded phrases? Without such operational definitions, researchers are unable to calculate values for the measures or evaluate their linguistic adequacy.

References

Biber, D., Gray, B., Staples, S., & Egbert, J. (2022). *The register-functional approach to grammatical complexity: Theoretical foundation, descriptive research findings, applications*. Routledge. https://doi.org/10.4324/9781003087991-24

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (2021). *The grammar of spoken and written English*. John Benjamins. (Previously published as *The Longman grammar of spoken and written English*, 1999). https://doi.org/10.1075/z.232

Biber, D., Larsson, T., & Hancock, G. R. (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. https://doi.org/10.1515/cllt-2023-0016

Biber, D., Larsson, T., & Hancock, G. R. (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. https://doi.org/10.1177/00754242231222

Biber, D., Larsson, T., Hancock, G. R., Reppen, R., Staples, S., & Gray, B. (2024). Comparing theory-based models of grammatical complexity in student writing. *International Journal of Learner Corpus Research*. Advance online publication. https://doi.org/10.1075/ijlcr.23036.bib

Huddleston, R., & Pullum, G. K. (2002). *The Cambridge grammar of the English language*. Cambridge University Press. https://doi.org/10.1017/9781316423530

Hunt, K. W. (1965). *Grammatical structures written at three grade levels*. National Council of Teachers of English.

Labov, W. (1969). Contraction, deletion, and inherent variability of the English copula. *Language*, *45*(4), 715–762. https://doi.org/10.2307/412333

Nichols, J. (1984). Functional theories of grammar. *Annual Review of Anthropology*, *13*, 97–117. https://doi.org/10.1146/annurev.an.13.100184.000525

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. Longman. https://doi.org/10.1177/007542428702000108

Szmrecsanyi, B. 2017. Variationist sociolinguistics and corpus-based variationist linguistics: Overlap and cross-pollination potential. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, *62*(4), 1–17. https://doi.org/10.1017/cnj.2017.34