EnglishUSA Journal

Volume 9 *Fall/Winter* Article 5

2024

The Purposes of First-Year Course Syllabi According to Corpus Data

Mark Makino
Southern Utah University, markmakino@suu.edu

Recommended Citation
Makino, Mark (2024) "The Purposes of First-Year Course Syllabi According to Corpus Data," *EnglishUSA Journal*: Vol. 9, Article 5.
DOI: http://doi.org/10.14305/jn.2836581X.2024.9.1.1.02
Available at: https://surface.syr.edu/englishusa_journal/vol9/iss1/5

The Purposes of First-Year Course Syllabi According to Corpus Data

Keywords

EnglishUSA, syllabus, concordancer, English language teaching, intensive English programs, applied linguistics, international education

Cover Page Footnote

The author collected 19 syllabi for first-year undergraduate courses at a mid-sized public university and analyzed them with the use of a concordancer. The analysis revealed strong tendencies for particular items to be overrepresented in syllabi compared to other genres of writing. The results also support a view of syllabi firstly as a set of expectations on both the student and instructor.

The Purposes of First-Year Course Syllabi According to Corpus Data by Mark Makino

Introduction

A meme circulated among academics on social media in which a university professor described a joke/experiment that he played on his students: he hid in his syllabus directions to open a specific locker on campus which, though unmentioned in the directions, contained a fifty-dollar bill (Wilson, 2021). The fifty-dollar bill remained unclaimed at the end of the semester, meaning that no student read the syllabus closely enough to notice the directions to the hidden treasure. The meme became national news (Smart, 2021) due to its whimsical illustration of the failure of university students to read assumedly important course syllabi. Although syllabi are certainly documents that some instructors spend significant time on and hope that students read carefully, the syllabi themselves may be less readable than is assumed. Indeed, they may pose special difficulty to international students and other English learners, both because of the language and because of their special purposes as documents.

Syllabi are a specialized type of academic writing with different purposes for different audiences. As Slattery and Carlson (2005) point out, syllabi serve a variety of functions for the student, the instructor, and the university administration, including supplying and recording relevant information about the course, recommending study habits, and providing a permanent record of pedagogical practices and topics covered. Parkes and Harris (2002) categorize the purposes of a syllabus into "the syllabus as a contract, the syllabus as a permanent record, and the syllabus as a learning tool" (p. 55), noting as Slattery and Carlson do that a syllabus is designed to contain different types of information for different audiences. There is the potential for a wide variety in content, although the purposes above provide areas of likely commonality across different syllabi.

Many faculty who work with international students have a stake in whether or not they can read this extremely common type of document, whose language and purpose may not be transparent. The genesis for this research was simply noticing that every university student is handed a small stack of these rather dense and legalistic documents at the start of every semester, and that my students, eventual graduates of the Intensive English Program (IEP) hosted at the university where I am employed, will be expected to read them. Whether syllabi follow similar patterns to other subgenres of academic English or have their own idiosyncrasies that we might do better to anticipate and prepare our students for spurred me to begin collecting syllabi for analysis.

Because higher-order features of syllabi have been analyzed in depth previously (for example, Parkes and Harris, 2002; Slattery and Carlson, 2005), I have decided to analyze lower-order linguistic features of the collected syllabi as a corpus with the use of a concordancer and put these in the context provided by the higher-order features.

Method

19 syllabi for first-year courses were collected, edited for formatting, edited for content, and assembled into a corpus for analysis, totalling 23,767 tokens. The syllabi were collected by visiting individual faculty offices at the mid-sized public university in the Western United States where I am also employed, explaining this research project, and obtaining syllabi

either on paper or by email. Of the 19 syllabi, 11 are from STEM fields and 8 are from the humanities, giving some balance to the different fields represented in the corpus. This was deliberate, as "Balance, representativeness and comparability are ideals which corpus builders strive for but rarely, if even, attain" (McNery & Hardie, 2011, p. 10). I also restricted the corpus to syllabi for courses designated for first-year students because these are the documents that IEP graduates or new international students would be expected to be able to read during their first semesters in degree programs.

Each syllabus also had specific sections removed in order to reduce redundancy in the corpus. The following sections were always omitted:

- Disability resources and accommodations
- Hyperlinks to required language
- University statements
- Department statements
- Academic honesty and plagiarism statements
- Emergency management statements
- Higher Education Opportunity Act (HEOA) compliance statements
- COVID-19 statements (dealing with masks, changing modalities, etc.)
- Zoom etiquette statements
- Course schedules

All of these beside the last were omitted simply because they were usually identical, and including them would have crowded out similarities that were due to the unique status and purposes of syllabi rather than the boilerplate requirements of the university. The last, course schedules, were omitted because they usually consisted only of dates and technical words specific to those courses.

The corpus was analyzed using a concordancer written by me for the frequencies of words from the Academic Word List (AWL) (Coxhead, 2000), as well as multi-word chunks (Nattinger & DeCarrico, 1992). Results of this concordancer were verified against results from AntConc (Anthony, 2022). These frequencies were compared to the frequencies of the same words and chunks in the Corpus of Contemporary American English (COCA) (Davies, 2008-). Within COCA, the academic genre specifically (henceforth COCA ACAD) as well as the corpus as a whole (henceforth COCA general) were used for comparison. The reason for this was to provide a rough baseline level of frequencies in academic English and for American English more generally, in order to determine whether the patterns observed in syllabi were unusual in comparison to other common genres of spoken or written English.

I compiled two separate lists of lexis from the AWL for closer examination: high frequency words of any word class (with at least 20 instances in the syllabus corpus) and verbs with their conjugations. I also compiled a list of high-frequency chunks which were not limited to words appearing in the AWL. The high frequency items were selected in order to compare their frequencies in the syllabi with their frequencies in academic writing generally as well as non-academic English. The verbs were chosen randomly except that each occurs in the corpus at least once. The reason for comparing verb conjugations is that the frequencies of those conjugations may inform the contexts in which English learners are familiar with seeing them and the contexts in which English instructors present them. For example, an instructor might be more likely to give an example sentence in which "require" is used in its past participle form in a sentence such as "X is required" as opposed to the preterite "The school required X". Thus, the results of analysis of the frequencies of different conjugations could inform future instruction. The results from concordancer searches for the above tokens were entered into a spreadsheet along with the results from searches for the same items on COCA, giving a

chart of the relative frequencies of a variety of tokens across three corpora (the syllabus corpus, COCA ACAD, and COCA general).

Results

All Words

The most common tokens overall in the syllabus corpus are also in the AWL can be seen below. Numerical results for each chart can be found in Appendix A. All frequencies are given per million tokens: first for the syllabus corpus, then COCA ACAD, and last COCA general.

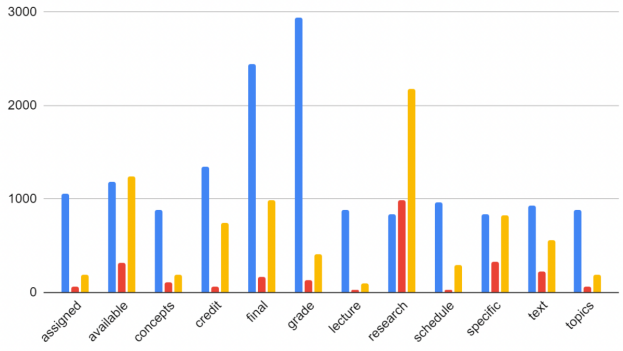

Figure 1. Frequent tokens in the syllabus corpus.

Among the tokens with at least 20 instances in the corpus, a pattern of overrepresentation in syllabi can be seen. The first is that words that one might expect to be more common in a syllabus are indeed much more common than in either academic English or American English overall. "Credit", for example, appears 1346.40 times per million words in the syllabus corpus, but only 62.50 times per million in COCA ACAD and 746.00 times per million in COCA general, where, as one might anticipate, it usually means something other than "academic units". However, some words that have a looser connection to course syllabi show a similar pattern. "Concepts", for example, occurs 883.58 times per million in the syllabus corpus (and in 10 of the 19 syllabi), 104.00 times in COCA ACAD and 184.79 in COCA general. Besides "research", all of the most common tokens in the syllabus corpus seem to be far more common in syllabi than in academic writing generally.

High-frequency tokens also have a high likelihood of appearing in syllabi at least once per document, and when they appear once, often appear multiple times. "Credit", for example, appears in 12 of the 19 syllabi, and in those 12 syllabi appears an average of 2.6 times. "Topics" appears in 10 of the 19 syllabi an average of 2.1 times. "Grade" appears in 17 syllabi 4.11 times per syllabus. Students are very likely to encounter the high-frequency tokens at least once per syllabus, and students unfamiliar with them will probably face greater difficulty in understanding syllabi

Verb Forms

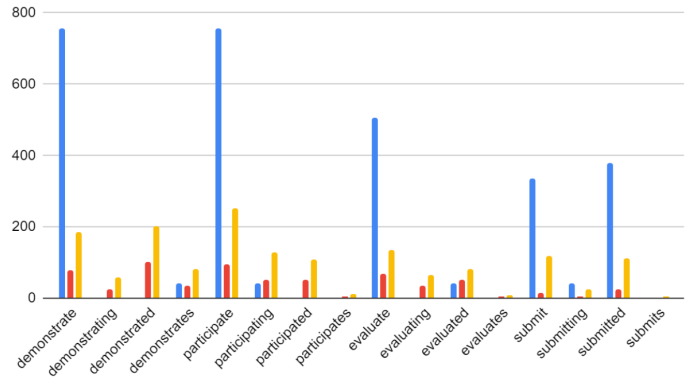

Figure 2. Frequent verbs in the syllabus corpus with conjugations.

For verb conjugations, the distributions also reveal a strong tendency for particular conjugations to appear far more frequently than the other conjugations do, and also for those conjugations to appear far more frequently in syllabi than in other genres of English. The chart above shows the frequencies of the forms of four verbs, and shows that typically, one form predominates in the syllabus corpus, while forms are more evenly distributed in the other corpora. The form that predominates in the syllabus corpus also doubles or triples the frequencies of the

same verbs in those forms in the other corpora.

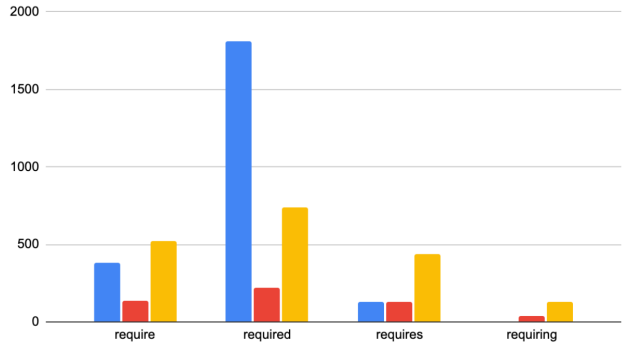

Figure 3. Conjugations of "require" across the three corpora (syllabi, COCA ACAD, and COCA general).

To zoom in on one particular verb, "require" displays the characteristics described above. In the syllabus corpus the past tense/past participle form "required" represents 78.18% of all instances of that verb. The distribution of the conjugations of "require" in COCA ACAD is more even: 25.9% for "require", 42.2% for "required", 24.4% for "requires" and 7.5% for "requiring". The total number of instances for the most common form is also more than eight times higher in the syllabus corpus than in COCA ACAD. It should be noted that my self-coded concordancer does not distinguish between past participles, past tenses, and participial adjectives, but upon reviewing the syllabi manually, all of the instances of "required" were past participles (e.g. "all students are required to…") or participial adjectives ("complete the required tasks"). COCA is not perfectly reliable in distinguishing these three forms, but according to its estimates, the past participle "required" occurs 171.4 times per million words, the adjective 40.1 per million, and the past tense 9.9 per million. The distribution of these forms in the syllabus corpus is, in other words, an exaggerated version of the same contour in COCA ACAD. The overwhelming predominance of one form matches the pattern seen for other verbs.

A further dimension of the frequency of particular words in syllabi is the likelihood that one will encounter them at least once in every single syllabus. The verb "require" in some form occurs 45 times total in the syllabus corpus in various forms, at least once in 18 of the 19 syllabi. In 13 of the syllabi it occurs as a past participle or participial adjective "required". Not only are students very likely to encounter this word in this form a certain number of times per million words, but they have a near-certain likelihood of encountering it at least once in each course syllabus that they receive. Understanding syllabi, therefore, depends disproportionately on understanding this one verb in one particular form.

Chunks

Chunks show some of the same tendencies as individual words; they are both much more common in syllabi than in other genres, and much more common than other apparently similar multi-word units.

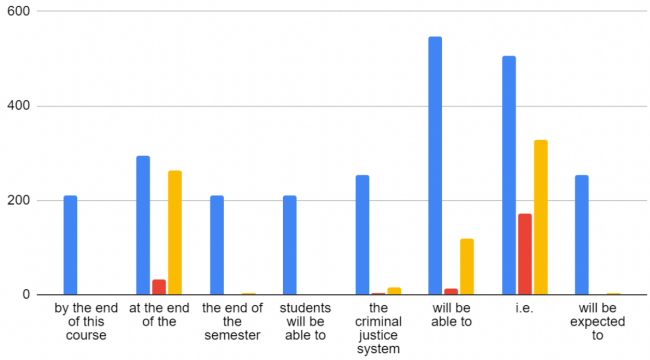

Figure 4. Frequent chunks of any size in the syllabus corpus

In the above list of the most frequent chunks of four to six tokens (note that punctuation counts as a token both in my concordancer and in COCA), one can see that both clearly syllabus-specific chunks and less obviously syllabus-specific chunks may predominate on syllabi. "The criminal justice system" is clearly an example of the former, occurring all of its six times in only one syllabus. "Will be able to" is an example of the latter, occurring more than 40 times as often per million in the syllabus corpus as in COCA ACAD and almost

five times as often as in COCA general, despite being composed of four common words (none of which are in the AWL) which on their face have little to do with syllabi in particular. Note that an apparently similar chunk like "should be able to" occurs only once in the syllabus corpus, and "may be able to" does not occur at all. The distribution of the chunk "will be able to" across syllabi is also noteworthy: it occurs in 7 of the 19 syllabi (36.8%), but in those 7, it occurs an average of 1.85 times each. In other words, when it appears in syllabi at all, it is likely to occur more than once. "Will be expected to" shows a similar pattern, occurring in 4 syllabi an average of 1.5 times each. Rather than being distributed evenly among almost all syllabi like the past participle of "require", chunks seem to be concentrated in fewer syllabi, but being repeated in those syllabi.

Discussion

The commonness of a particular word or chunk on course syllabi does not necessarily indicate a need to prepare English learners specifically for it. However, the prevalence of the specific words and chunks highlighted in the previous section supports a particular purpose of syllabi that students may need to be aware of, and also support the need to interpret those words and chunks in particular ways which could benefit from instruction.

The frequent words and chunks highlight the role of the syllabus as a contract (Parkes and Harris, 2002). "Like any contract, the syllabus serves to set forth what is expected during the term of the contract—typically a semester—and to guide the behaviors of both parties" (p. 55). This role of the syllabus casts words such as "required" and chunks such as "will be expected to" in a new light. Seen through Speech Act Theory (Searle, 1976), these would best be considered directives, "attempts… by the speaker to get the hearer to do something" (p. 11). In another type of document, a sentence from the syllabus corpus such as "Students will be expected to attend on that day in order to receive full credit" could be interpreted simply as *representative*, or "to commit the speaker … to the truth of the expressed proposition" (p. 10), as with the usage earlier in this article, "[Students] at the university where I am employed, will be expected to read them". Given that one purpose of the syllabus is to direct students to behave in particular ways, certain words or chunks may be used for that purpose, although students may be familiar with them for different purposes. Because students may not interpret them as directives without an understanding of the syllabus as a contract, specific instruction on context-specific interpretations may be warranted.

The same purpose of the syllabus as a contract may be said for some expressions which implicitly or explicitly commit the instructor to some actions. Searle's (1976) category of *commissives* seems to apply to utterances from the syllabus corpus such as "All errors will be corrected but must be brought to our attention within one week of the papers being returned", serving as a directive to students but as a *commissive* to a the instructor. The modal auxiliary "will" combined with the passive voice "be corrected" here serves "to commit the speaker (again in varying degrees) to some future course of action" (Searle, 1976, p. 11). Needless to say, this *commissive* interpretation of "will" may also not be obvious to students even if they are familiar with other usages.

The other proposed functions of syllabi, *syllabus as permanent record* and *syllabus as a learning tool*, are suggested less than *syllabus as a contract*. One also sees reminders of the syllabus as permanent record in the prevalence of certain words and chunks "providing details of what was covered, what students were expected to do, and how these outcomes and performances were assessed" (Parkes and Harris, 2002, p. 57). These words

and chunks may be interpreted in particular ways through this lens. Chunks such as "will be able to" often describe desired outcomes as part of an official record for the course, for example in the phrase "Therefore, by the end of this course, students will be able to…". The prevalence of particular chunks support interpretation of syllabi mostly as *syllabus as a contract* or *syllabus as permanent record*.

Besides indicating and being interpreted in the light of some specific purposes of syllabi, the words and chunks sometimes indicate bottleneck-like function of a small number of words and chunks; that is, a small number of key lexical items may facilitate or prevent comprehension of the document. The bottleneck effect that these items have is not due just to their commonness per million words but their presence at least once, and usually more than once, in nearly every syllabus.

Conclusions

The syllabi reviewed for this article feature repeated use of a limited number of words, verb forms, and chunks compared to other types of academic writing or other genres of English not limited to academic writing. An English learner is very likely to need these words and chunks to even minimally comprehend course syllabi that they encounter in their first year of college. Consideration of the syllabus as a special kind of document will likely help them to interpret these words and chunks correctly according to the intentions of their writers.

References

Anthony, Laurence. (2022). AntConc (Version 8.5.15) [Computer Software]

Biber, D. (2006). University language (Vol. 10). Amsterdam: John Benjamins.

Charles, Maggie. (2015). Genre, Corpus, and Discourse: Enriching EAP Pedagogy. In: Thompson, P., & Diani, G. (Eds.). (2015). English for Academic Purposes: Approaches and Implications. Cambridge Scholars Publishing.
Coxhead, A. (2000). A new academic word list. TESOL quarterly, 34(2), 213-238.

Coxhead, A. (2000). A new academic word list. TESOL quarterly, 34(2), 213-238.

Davies, Mark. (2008-) The Corpus of Contemporary American English (COCA). Available online at https://www.english-corpora.org/coca/.

McEnery, T., & Hardie, A. (2011). Corpus linguistics: Method, theory and practice. Cambridge University Press.

Nattinger, J. R., & DeCarrico, J. S. (1992). Lexical phrases and language teaching. Oxford University Press.

Palumbo, Giuseppe. Studying ELF Institutional Web-based Communication by Universities: Comparison and Contrast with English Native Texts. In: Thompson, P., & Diani, G. (Eds.). (2015). English for Academic Purposes: Approaches and Implications. Cambridge Scholars Publishing.

Parkes, J., & Harris, M. B. (2002). The purposes of a syllabus. College Teaching, 50(2), 55-61.

Slattery, J. M., & Carlson, J. F. (2005). Preparing an effective syllabus: Current best practices. College Teaching, 53(4), 159-164.

Searle, J. R. (1976). A classification of illocutionary acts. Language in society, 5(1), 1-23.

Smart, S. (2021, December 18). A professor hid a cash prize on campus. All students had to do was read the syllabus. CNN.com. https://www.cnn.com/2021/12/18/us/tennessee-professor-syllabus-money-trnd/index.html

Thompson, Paul. (2015). Changing the Bases for Academic Word Lists. In: Thompson, P., & Diani, G. (Eds.). (2015). English for Academic Purposes: Approaches and Implications. Cambridge Scholars Publishing.

Van Rooy, B., & Terblanche, L. (2009, January). A multi-dimensional analysis of a learner corpus. In Corpus Linguistics (pp. 239-254). Brill.

Wilson, Kenyon. (2021, December 8). My semester-long experiment has come to an end. At the start of the term, I placed $50 in one of… [Image attached] [Status update]. Facebook. https://www.facebook.com/kenyon.wilson.7712/posts/1480005132394051

Mark Makino grew up in Orange County, California, before moving to Japan and much later to Cedar City, Utah, which is much more similar to Japan than it is to Orange County.

Appendix A:

	SYL PM	COCA ACAD PM	COCA PM
assigned	1051.88	61.80	185.00
available	1178.10	313.30	1239.48
concepts	883.58	104.00	184.79
credit	1346.40	62.50	746.00
final	2440.36	159.10	987.03
grade	2945.26	135.00	403.93
lecture	883.58	26.40	99.98
research	841.50	983.80	2179.99
schedule	967.73	29.40	292.42
specific	841.50	321.60	829.01
text	925.65	219.50	560.49
topics	883.58	61.70	182.48

	SYL PM	COCA ACAD PM	COCA PM
demonstrate	757.35	78.60	183.92
demonstrating	0.00	23.40	57.83
demonstrated	0.00	100.10	202.23
demonstrates	42.08	34.40	81.90
participate	757.35	94.70	252.23
participating	42.08	52.80	128.72
participated	0.00	50.30	109.77
participates	0.00	5.00	13.08
evaluate	504.90	68.50	133.97
evaluating	0.00	35.90	66.41
evaluated	42.08	50.20	82.52
evaluates	0.00	4.20	8.89
submit	336.60	16.00	119.80
submitting	42.08	4.40	24.94
submitted	378.68	23.80	110.07
submits	0.00	1.30	6.08
require	378.68	135.7	521.23
required	1809.23	221.3	739.28
requires	126.23	128.2	437.65
requiring	0.00	39.7	131.04

	SYL PM	COCA ACAD PM	COCA PM
by the end of this course	210.38	0.00	0.01
at the end of the	294.53	32.00	262.14
the end of the semester	210.38	1.40	2.90
students will be able to	210.38	0.80	1.36
the criminal justice system	252.45	2.60	14.42
will be able to	546.98	12.40	118.27
i.e.	504.90	171.80	327.66
will be expected to	252.45	0.50	2.71
if you are	420.75	9.30	333.79
there will be	631.13	14.30	264.88
the due date	546.98	0.30	2.85
will not be	631.13	21.00	206.73
this course will	462.83	0.00	0.31
in this course	673.20	0.80	1.33
the final exam	504.90	0.40	1.40