Nominalization in high- and low-rated L2 undergraduate writing

Tetyana (Tanya) Bychkovska (independent scholar)
Joseph J. Lee*, Dalarna University, Sweden
(*Corresponding Author: jole@du.se)

Nominalizations, or nouns derived from verbs or adjectives through suffixes, are a pervasive characteristic feature of written academic discourse. To better understand the nature of nominalization in L2 student writing and its relation to assessment in first-year writing (FYW) contexts, we report findings of a comparative corpus-based analysis of nominalization use in university student papers. Data consist of high-rated (A graded) and low-rated (C graded) L2 undergraduate research papers from multiple sections of an FYW course for international and multilingual students. Nominalizations were examined in terms of frequencies, unique types, abstract/concrete and human/non-human categories, nominal stance types, and modification types. Results reveal no statistically significant differences in the examined classifications. However, the small effect sizes for certain categories point to subtle differences between the two groups, which together might have affected the instructors' evaluations of text quality. We conclude with suggestions for incorporating nominalization instruction in English for Academic Purposes writing courses.

Keywords: nominalization, nominal features, second language writing, academic writing, corpus analysis

Introduction

In English for Academic Purposes (EAP) writing research, nominal features have received increasing attention over the past decade. Such attention is not surprising since corpus-based investigations have demonstrated that approximately two-thirds of the content words in academic prose are nouns (Biber et al., 1999). Research has also shown that the frequency of nominals in academic texts has increased over the past three centuries (e.g., Banks, 2008; Biber & Gray, 2013). Biber et al. (2013) argue that academic writing is characterized by its 'reliance on phrasal structures, especially complex

https://doi.org/10.3828/ijeap.2023.8
ISSN 2634-4610 ONLINE

phrases with phrasal modifiers' (p. 192). Thus, noun phrase features have been proposed to be a more accurate measure of grammatical complexity in academic prose than the dependent clause-based measures traditionally used (Biber et al., 2011). These findings suggest that novice academic writers can benefit from learning about how to construct prose reliant on nominal features. One common strategy used to convert clauses into phrases without considerably altering the original meaning is called nominalization, or changing the main verb to a noun (e.g., *argue – argument*) or *be*-copula plus adjective into a noun (e.g., *be important – importance*). Nominalization permits writers to express complex ideas concisely, to shift the focus from human agents to processes and concepts, and to refer to previously mentioned information showing cohesive ties (Biber et al., 1999; Halliday, 1989; Liardét, 2013). Considering that these functions are crucial for successful academic writing (Biber & Gray, 2013, 2016), gaining control over this strategy can lead novice EAP writers to produce texts that effectively meet the expectations of this register's readers.

Nominalization in academic writing

Nominalization, or the process of creating nouns from adjectives or verbs, is a characteristic feature of academic writing. As Biber and Gray (2016) report, it is used more frequently in academic prose than in other written registers such as newspapers and novels. Nominalization is commonly used in academic writing because of the essential functions it serves in such texts. First, by removing the agent of the action, nominalization allows writers to shift the focus from human participants to abstract objects that carry the meaning of actions or processes (Biber et al., 1999). This human-to-object shift adds to a text's formality and objectivity. To illustrate this function, Baratta (2010, p. 1019) offers the following examples:

1) I discovered similar findings on this subject during the research.
2) The discovery of similar findings on this subject during the research...

As can be seen, *I discovered* (1) is converted into *discovery* (2). Deleting the agent and placing the emphasis on *the discovery*, according to Baratta, 'allows for more objectivity' in the proposition (p. 1019).

Another function of nominalization is information compression, a key characteristic of academic writing. Nominalization assists in word economy and meaning 'packing', thus leading to concise texts rich in meaning (Biber, 1988). In 2) above, nominalization allows the writer to pack additional information within a sentence, unlike in 1) where the writer would need to add information in a new sentence. Condensing potentially several sentences into one contributes to both lexical density and text concision.

Lastly, nominalization serves to achieve greater text consistency and

cohesion. For example, Norman (2003) examined twelve possible lexico-grammatical realizations of anaphoric references to key terms in biomedical research article abstracts. After repetition, nominalization was found to be the most frequent device used to maintain consistency and cohesion in these abstracts. Nominalization may serve as either anaphoric or cataphoric reference, allowing writers to refer to a proposition mentioned in the previous or following discourse to maintain connections between ideas.

To better understand the role that nominalization plays in academic writing, EAP researchers have examined it from different perspectives. Some studies (e.g., Baratta, 2010; Yoon, 2018) have viewed nominalization as a derivational process of creating nouns from verbs or adjectives by adding suffixes (e.g., *place – placement*, *happy – happiness*), with the most productive suffixes being the following: *-ment*, *-ance*, *-tion*, *-ness*, and *-ity* (Biber, 1988). Nominalization has also been investigated from the perspective of systemic functional linguistics (Halliday, 1989) as one of the realizations of grammatical metaphor (e.g., Liardét, 2013; 2016; Ryshina-Pankova, 2010). Grammatical metaphor in this case occurs when a noun, which typically carries the meaning of 'entity', takes on the meaning of 'process' or 'quality', normally expressed by verbs or adjectives, respectively.

Researchers from these two orientations use different methodological approaches in their examination of nominalization. Studies adopting the grammatical metaphor approach, which identifies all nouns carrying the meaning of verbs and adjectives, tend to examine a smaller number of texts using qualitative methods. Research in the derivational perspective, conversely, employs a corpus-based approach focusing on nominalization in a large number of texts through quantitative methods, which can lead to greater generalizability. Corpus-based studies, however, have often limited their analysis to nominalizations formed with only 5–7 suffixes. While these suffixes are among the most productive (Biber et al., 1999), limiting the number of suffixes could lead to an incomplete picture of nominalization usage in academic writing. Further, corpus-based studies promote our understanding of quantitative uses of nominalization, but they seldom provide more detailed qualitative analysis of its use. Little corpus-based research has investigated, for example, the functional aspect of nominalizations (e.g., stance), their semantic categories, and the noun phrase structures (i.e., modification). Examining these dimensions, however, would contribute to closing the form-function gap and provide a more comprehensive understanding of how nominalization is used in academic writing.

Nominalization in L2 student writing for academic purposes

As a key feature of academic writing, nominalization has been commonly included in research on lexico-grammatical features in university students' texts. Use of nominalization has been found to distinguish student writing

at different academic levels. In his longitudinal case study in the U.K., Baratta (2010) examined how nominalization was used by the same six students (first-language [L1] backgrounds unspecified) over three years of undergraduate study. He found that the average nominalization use increased minimally from year one to two, but the increase was prominent from the second to the third year of study. He attributes these differences to the text types (e.g., personal reflections vs research papers) that students produced. In Staples et al'.s (2016) cross-sectional study of L1 English student writing across four university levels (undergraduate years one to three and graduate) in the U.K., it was found that nominalization use increased with students' educational experience. Similar to Baratta (2010), they also found variation in the use of nominalization across genres as well as disciplines in all groups. In addition to demonstrating that the frequency of nominalization increases with students' educational level, studies show that genre, register, and discipline are important variables to consider in examining nominalization use in student writing (Baratta, 2010; Gardner et al., 2019; Staples et al., 2016).

Studies on second-language (L2) student writing have also shown that the frequency of nominalization increases with experience. Crosthwaite (2016), for example, found that nominalization occurred more frequently in L2 student writers' texts after receiving training in an EAP course for only one term. Liardét (2013; 2016) also identified an increased use of nominalization in L2 student texts as the learners in her research progressed through their programme of study. Her analysis, however, revealed that their use was frequently characterized by ambiguity (Liardét, 2013) and did not necessarily contribute to better cohesion in writing (Liardét, 2016). These results suggest that, although L2 students employ a greater frequency of nominalization as they gain experience and proficiency, they may still face challenges in expressing their ideas cohesively through nominalization.

When L2 student essays were compared to L1 texts, however, few quantitative differences were found. In Hinkel's (1997) examination of nominalization, along with twenty other features, in 150 placement tests written by L1 and L2 students, no difference was identified in nominalization usage. Similarly, in his study of nominalization use in L1 English and Korean English-as-a-foreign-language student writers, Yoon (2018) also found no differences in general frequency of this feature between the two groups; however, his more detailed analysis revealed that L2 writers modified the nouns (i.e., pre- or postmodification) less frequently than the L1 students.

The use of nominalization has also been studied in relation to writing quality. In their comparison of three proficiency levels of *Test of Written English* (*TWE)* essays, Grant and Ginther (2000) found that nominalization was one of the main predictors of writing proficiency, with higher-proficiency

writers using nearly twice as many nominalizations. Similar to most studies on undergraduate L2 writing, however, their study examined short-timed essays, the results of which may be limited in generalizing to the types of writing that university students are expected to produce for their study (Leedham, 2015).

While these studies of student writing have revealed that the frequency of nominalization increases with writers' academic standing (Staples et al., 2016), progress in an EAP course (Crosthwaite, 2016), and writing proficiency (Grant & Ginther, 2000), the majority of this research includes nominalization as one among other linguistic features under investigation. Few studies have exclusively examined nominalization; thus, they have provided limited analysis of nominalization beyond general frequencies. Furthermore, studies that have specifically focused on nominalization use in student writing have investigated the longitudinal development among only a few writers (e.g., Baratta, 2010; Liardét, 2013), or the analysis has been based on short-timed texts in response to prompts (e.g., Liardét, 2016; Yoon, 2018), 'requiring no research or preparatory reading' (Leedham, 2015, p. 33). Restricted in scope and highly decontextualized, such writing tasks do not represent the kinds of writing expected of university students and reveal little about how L2 students use nominalization in source-based research papers typically required of them. These prompt-based writing tasks are influenced by prompts, 'task setting (time availability) and intertextuality (access to secondary sources)' (Ädel, 2008, p. 35).

In addition, although research has revealed differences between general English abilities and academic writing proficiency, we know little about the use of nominalization in source-based research writing by L2 university students in the context of first-year writing (FYW), or the relationship between nominalization use and assessment of L2 FYW student writing. To better understand the nature of nominalization and its role in L2 academic writing development, we report findings of a comparative corpus-based analysis of the types and functions of nominalization in high- and low-rated L2 university student research papers. This study aims to contribute to our understanding of how English-as-a-second-language (ESL) students' grammatical choices relate to their writing assessment within the context of FYW. The study was guided by the following research questions:

1. To what extent do high- and low-rated L2 research papers differ in their frequency of use of nominalization?
2. To what extent do high- and low-rated L2 research papers differ in the semantic and structural types of nominalization?
3. To what extent do high- and low-rated L2 research papers differ in the stance options of nominalization?

Materials and methods

Description of the corpora

The data consist of 133 assessed research papers written by ESL undergraduate students in a U.S.-based FYW course. These papers are derived from the subset of the Corpus of Ohio Learner and Teacher English (COLTE), a large collection of a range of assessed L2 student texts, ranging in grades from A to F, and teachers' electronic written feedback at Ohio University. The student papers selected for the present study come from multiple sections of the second course in the FYW sequence across several semesters. The FYW course is a specifically designed EAP course for international and multilingual undergraduate students and taught by ESL writing specialists with a minimum of an MA in TESOL/applied linguistics. While some students placed into this FYW course earn a TOEFL iBT writing section score of twenty-four or higher, the majority enrol in this course after completing the first FYW course with a grade of C or higher.

The standardized curriculum develops L2 students' higher-level academic writing abilities to be successful in disciplinary courses: compose effective papers for different purposes and audiences; analyse purposes and audience for various academic genres; engage in secondary research; integrate sources through paraphrasing, summarizing, and quoting, following APA style; use appropriate academic grammar and vocabulary; and self-edit for grammatical accuracy. At the time the corpus was compiled, nominalization instruction was not included in either of the courses in the FYW sequence. Successful completion of this course fulfils students' institutional FYW requirements for graduation.

The research papers selected were written by sixty-seven female and sixty-six male ESL students from a variety of L1 backgrounds: Mandarin (n = 86), Arabic (n = 26), Korean (n = 3), Spanish (n = 3), Cantonese (n = 2), Finnish (n = 2), Greek (n = 2), Thai (n = 2), and one from eight other languages (Armenian, Brazilian Portuguese, Gujarati, French, Japanese, Macedonian, Norwegian, and Twi). This distribution of L1 backgrounds is typical of the student population in this FYW course. Students' average age was 21.6 years (*SD* = 1.8), and they, on average, studied English for 7.8 years (*SD* = 3.4) in their home countries and 4.1 terms (*SD* = 1.8) in U.S.-based intensive English programmes.

The source-based research paper assignment required students to choose and research specific issues related to broader course themes (e.g., natural disasters, refugee crisis, public health, media bias). As the papers were drawn from a corpus of naturalistic classroom written assignments, the students wrote on a variety of topics within the broader themes of the course. The corpus includes papers submitted for both a shorter, four-page assignment, and a longer, eight-page assignment. In both tasks, students were expected

to analyse and synthesize source content and views and use at least four academic sources for the shorter task and a minimum of six academic sources for the longer task. The assignments required students to develop research questions, conduct extensive library research using both primary and secondary sources, and analyse and evaluate the sources in their papers. Students submitted an outline and two or three drafts, depending on the course instructor. Instructors graded and provided electronic written feedback on each draft following a standardized grading rubric that all teachers of this FYW course used. The grading rubric consists of categories of content (40%), organization (30%), source use (10%), and language use (20%). Final drafts were selected for analysis in the present study.

Each research paper was manually cleaned: we removed the paper codes, titles, section headers, footers, figures, appendices, and references. Upon cleaning the papers, the total word count of the corpus was 192,407 words. The data consist of two groups based on the grades the papers were assigned by course instructors: high-rated (A papers; $M = 94.21\%$, $SD = 2.63$) and low-rated (C papers; $M = 74.65\%$, $SD = 3.36$) groups. Table 1 presents the descriptive statistics for the high-rated corpus (HC) and low-rated corpus (LC).

Table 1 Description of the two corpora

	HC	LC
N	84	49
Word total	127,609	64,798
Word mean	1,519	1,322
SD	617	486

Note: *SD* = standard deviation

Methods of analysis

Identification of nominalization

In this study, nominalization is operationalized as a range of nouns grammatically derived from verbs (e.g., *agree – agreement*) or adjectives (e.g., *different – difference*) through derivational suffixes. Gerunds, while included in some studies on nominalization (e.g., Baratta, 2010), were excluded from the present research because gerunds, as Liardét (2016) argues, are incomplete (i.e., developmental) realizations of nominalization.

Another consideration regarding the feature under analysis is related to 'faded metaphors'. Faded metaphors, the concept used in systemic functional linguistics studies, refer to nominalizations that once were metaphorical but have lost their meaning as a process because of frequent usage (e.g., *relationship*, *discussion*) (Derewianka, 2003). As they are not considered to be

metaphorical realizations, these types of nominalizations were excluded in a few studies (e.g., Yoon, 2018). Most researchers, however, have included faded metaphors because the decision of whether a certain nominalization is a faded metaphor seems to be dependent on researchers' intuition rather than on objective identification criteria (Liardét, 2013). Therefore, we included all nominalizations derived from verbs or adjectives into the analysis, irrespective of whether they could be considered faded metaphors.

To identify nominalizations, nineteen suffixes and their plural forms (see Appendix for the complete list) were searched in our corpus. We focus on a comprehensive number of suffixes unlike previous corpus research, which commonly used 5–7 suffixes to extract nominalizations. There are several possible reasons for this limited number of suffixes used in previous studies. The most commonly indicated rationale is that 5–7 suffixes are the most productive in English, forming the largest number of nominalizations in academic prose. However, it is worth noting that these most productive suffixes were identified based on research on published academic prose (Biber, 1999 et al., p. 323); it is possible, however, that for L2 undergraduate writing the most productive suffixes might be different. Another potential reason for relying on only several suffixes is that perhaps they can be used to retrieve nominalizations from a corpus with high accuracy (i.e., precision), eliminating the need for the manual check of all concordance lines. This reason is especially relevant for studies examining large corpora or having a wider scope (i.e., looking at many lexico-grammatical features or additional independent variables). Finally, in some cases, the choice to use a limited number of suffixes is arbitrary or unexplained. Since the focus of our study is narrow and is exclusively on nominalizations, we chose to search for nominalizations formed with all suffixes rather than to be selective and potentially miss important nominalizations in student writing. To compile a comprehensive list of nominalization-forming suffixes, we consulted previous nominalization research (e.g., Baratta, 2010; Yoon, 2018) as well as many grammar reference resources (e.g., Biber et al., 1999), until the search became exhaustive.

After retrieving all nominalization instances using AntConc (Anthony, 2018), we manually examined the concordance line and full textual context for each example. We excluded instances 1) if they contained letters from suffixes but were not nominalizations (e.g., *actu*al, *degr*ee); 2) if they were parts of entity names (e.g., *World Health Organization*); 3) if they were forms with no one-word equivalents (e.g., *filmmaker*); 4) if they were within quoted material; and 5) if they were used to modify a noun rather than being a head noun in the noun phrase (e.g., *violence outcome*, *employment crisis*). The last criterion was established because it was important to avoid conflating different syntactic functions that a nominalization might have in a noun phrase as this would pose additional challenges for the later modification

analysis. This exclusion criterion was also used by Yoon (2018), and in general, instances of a nominalization modifying a noun were rare in our corpus.

Analysis of nominalization

To address research question 1, we counted the number of nominalization tokens, types, and derivational suffixes in both corpora. Since previous research (e.g., Baratta, 2010; Grant & Ginther, 2000; Yoon, 2018) on nominalization typically focused on only 5–7 derivational suffixes, and this study expanded the list to nineteen, we also conducted a comparative analysis to determine whether focusing on a different set of suffixes would lead to different trends. For this purpose, in addition to counting the frequencies of nominalizations formed with nineteen suffixes, we also counted separately those formed with only five suffixes (i.e., *-ment*, *-ance*, *-tion*, *-ness*, *-ty*). These five suffixes were chosen because they were commonly used in previous research, and especially in the studies relying on the Biber Tagger (Biber, 1988), the part of speech tagger. Finally, we also searched top twenty most frequent nominalizations from each corpus in the Academic Word List (AWL, Coxhead, 2000) to identify whether these nominalizations are considered academic words.

Classification of nominalization

To address research questions 2 and 3, upon retrieving nominalizations, we classified them as 1) being abstract or concrete and human or non-human; 2) representing specific types of nominal stance; 3) and having pre- and/or postmodification. The categories were created after consulting previous research on nominalization (e.g., Liardét, 2016; Yoon, 2018) and nominal features (e.g., Jiang & Hyland, 2015).

Abstract/concrete and human/non-human nominalization. Identified nominalizations were classified as abstract if they referred to concepts having no physical representation in the real world (e.g., *awareness*, *legitimacy*) and as concrete if they referenced a person or tangible object that can be found in the real world (e.g., *director, printer*). Those nominalizations classified as concrete were further categorized as human, denoting people (e.g., *editor, employee*), or non-human, signifying objects or entities (e.g., *equipment*, *manufacturer*).

Nominalization stance. We also examined the stance options of abstract nominalizations, using Jiang and Hyland's (2015) functional taxonomy of stance nouns. Stance is defined as expressions of 'personal feelings, attitudes, value judgments or assessments' (Biber et al., 1999, p. 99), permitting writers to communicate their epistemic and affective position. In student writing, Lancaster (2014) argues that stance can 'play a role in shaping readers' judgments of overall writing quality' (p. 28). Jiang and Hyland's (2015) framework comprises three broad stance noun categories: *entity*, *attribute*,

and *relation* (see Table 2). *Entity* nouns express a writer's orientation to texts, events, discourses, or cognitive dimensions; *attribute* nouns convey a writer's assessment of the quality, status, and formation of entities; and *relation* nouns concern a writer's understanding of connections and relationships between entities. Based on this classification system, we coded each nominalization in functional terms.

Table 2 Stance noun framework (adapted from Jiang & Hyland, 2015, p. 535)

Entity	**Description**	**Examples**
Text	metatext	*paper*, *research*, *study*
Event	actions, processes, and real-world cases	*action*, *demonstration*, *observation*
Discourse	verbal expressions and speech acts	*argument*, *discussion*, *suggestion*
Cognition	attitudes, beliefs, and mental reasoning	*assumption*, *decision*, *expectation*
Attribute	**Description**	**Examples**
Quality	evaluative quality of traits	*importance*, *limitation*, *usefulness*
Manner	circumstances and formation of actions	*condition*, *direction*, *mechanism*
Status	epistemic, deontic, or dynamic modality	*ability*, *possibility*, *probability*
Relation	**Description**	**Examples**
Connection	relevance, cause-effect, difference	*difference*, *relation*, *relationship*

Nominalization modification. To understand the noun phrase structures of nominalizations, we classified them as containing premodification, postmodification, both pre- and postmodification, or no modification. Premodification included an attributive adjective (e.g., *inaccurate representation*, *written expression*); noun (e.g., *goods manufacturer*, *project description*); possessive noun (e.g., *director's decision*, *buyer's attention*); or a combination of premodifiers (e.g., *U.S.-China trade cooperation*, *potential biological complication*). Postmodification included a prepositional phrase (e.g., *portrayal of Muslims*, *proposal for possible solutions*); finite relative clause (e.g., *marriage that ensures…*, *students I interviewed*); non-finite relative clause (e.g., *penalties coming from…*, *students living on campus*); non-finite passive relative clause (e.g., *camper equipped with kitchen, convenience associated with…*); finite noun-controlled complement *that*-clause (e.g., *possibility that some of the elderly refugees lose…*); non-finite noun-controlled complement *to*-clause (e.g., *ability to manage*, *responsibility to provide*); or appositive (e.g., *ethnicities, especially ethnic minorities; the actor Jim Carrey*). If both premodification and postmodification of the nominalization were present, we counted them as one instance of the

'pre- and postmodification' category. In Yoon's (2018) analysis, possessive and demonstrative determiners were included; we, however, excluded these features because determiners are a required grammatical category rather than expansions of noun phrases.

Interrater agreement and statistical analysis

Each author worked independently to manually locate and analyse each nominalization in terms of abstract/concrete, human/non-human, modification, and stance. Interrater agreement was assessed with Cohen's kappa (κ) for each of the main categories investigated: κ =.96 for abstract/concrete, κ =.95 for human/non-human, κ =.98 for modification, and κ =.84 for stance. These kappa statistics indicate high interrater reliability. Remaining differences were discussed until we reached full agreement. Upon classifying all instances in the corpora, item frequencies were counted per text and normalized per 1,000 words (ptw). To determine whether the differences between the high and low groups were statistically significant, a series of Mann-Whitney U tests was conducted. This non-parametric alternative to the independent samples *t*-test was used since the majority of the data were not normally distributed. The alpha was set at.05 (two-tailed), and Bonferroni correction was applied for multiple comparisons. Effect sizes, Pearson's *r* values, were interpreted as small for ~.10, medium for ~.30, and large for ~.50.

Results and discussion

Comparison of nominalization tokens, types, and suffixes

Analysis of the overall frequency of nominalization formed with nineteen suffixes in the two corpora revealed no significant differences (see Table 3). Similarly, no significant differences were found when we examined nominalizations formed with only five suffixes. As shown in Table 3, however, focusing on a limited number of suffixes rather than the full range could reveal a slightly different trend. Although no significant differences were found for the five suffixes, there was a small effect (*r* =.15), suggesting that the two groups had some differences in the quantitative use of nominalizations. Therefore, it is important to focus on a comprehensive list of suffixes when examining nominalization in student corpora.

When these results were compared to a study that investigated the relation of nominalization to essay scores, similar overall patterns were found. Grant and Ginther (2000) who examined nominalizations formed from five suffixes in *TWE* essays found that higher-scoring essays included a greater number of nominalizations. While Liardét (2013, 2016) and Crosthwaite (2016) also found an increase in nominalization use as L2 English writers progress in their university education, these studies did not explicitly mention how

many suffixes were used. This methodological information, as supported by results in Table 3, is crucial for making accurate cross-study comparisons.

Table 3 Descriptive statistics and results of Mann-Whitney U for nominalizations in the two corpora

	HC		LC			
No. of suffixes	*Mdn*	*IQR*	*Mdn*	*IQR*	*p*-value	*r*
19 suffixes	38.44	21.14	35.03	21.90	.805	.02
5 suffixes	25.05	14.85	22.00	14.45	.078	.15

Note: Mdn = median normed frequencies per 1,000 words; *IQR* = interquartile range; statistical significance with Bonferroni correction is at $p<.017$.

Both corpora employed a full range of suffixes to form nominalizations. Similar to previous research (Biber et al., 1999), the most common suffix was *-(s/t)ion*, accounting for more than 30% of nominalizations in both corpora (see Figure 1). This suffix was followed by *-ty* (13%), *-er* (12%), and *-ment* (11%) in the HC, and by *-nt* (16%), *-er* (14%), and *-ty* (9%) in the LC. It is worth noting that all four of the most commonly used suffixes in the HC are the suffixes that form abstract nominalizations, while only two out of four most common suffixes in the LC were used to form this type of nominalizations (with the other two typically used to form concrete nominalizations). Interestingly, the suffixes *-nt*, *-th*, and *-or* were used proportionally more frequently than *-ness*, but they have not been typically included in nominalization research.

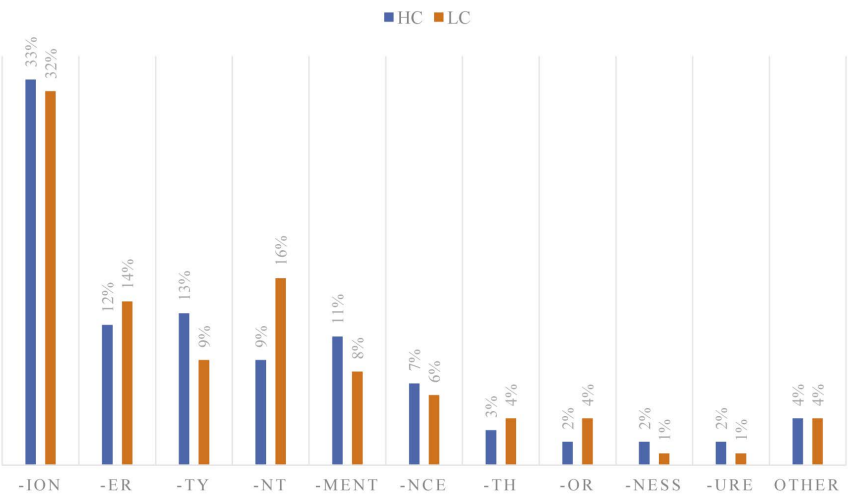

Figure 1 Distribution of derivational suffixes forming nominalization in the two corpora

Note: the OTHER category includes the suffixes -acy, -age, -al, -dom, -ee, -ism, -ist, -ncy, and -y.

In terms of unique types, operationalized as the number of nominalization types in each text, the HC employed only a slightly greater variety of nominalization types (*Mdn* = 22.00, *IQR* = 18.75) than the LC (*Mdn* = 21.00, *IQR* = 9.00), and the difference was not statistically significant (*p* =.18). Nevertheless, to some extent, the high-rated papers exhibited a wider nominalization repertoire than the low-rated papers, although the effect was small (*r* = .12).

This broader range in repertoire is further supported by the top twenty most commonly used nominalization types for each group, as shown in Table 4. For the LC, the twenty words comprise over 47% of the total instances of nominalization. In contrast, the top twenty words only account for approximately 36% of the total nominalization instances in the HC. In both corpora, *student*, *government*, and *information* are the top three words. Yet these words accounted for 14% of the total instances of nominalization in the HC, while they comprised 24% of the nominalization instances in the LC. *Student* alone accounted for 15% of all nominalizations used in the low group's writing, but only about 7% in the high group's texts. The frequent use of *student* in both groups can be explained by the writers' status as college students writing on issues relevant to their lives. Nominalizations, such as *government*, *violence*, *marriage*, *player*, *printer* and other specific nouns, are also connected to the topical matters chosen by the student writers. As previous studies have suggested (e.g., Grant & Ginther, 2000), there appears to be a relationship between topic and nominalization use. Yet as Thompson (2010) found in his analysis of high-rated dissertations (HRDs) and low-rated dissertations (LRDs) written by L2 MA students, 'topic-related nominalizations do not correlate with differences in rating – in fact, contrary to what might be expected, the LRDs have a higher average frequency than HRDs in this category' (p. 31). While the relationship between topic and nominalization is yet unclear, this is an area that may need further exploration (Grant & Ginther, 2000). The remaining nominalizations (e.g., *development*, *information*, *situation*, *solution*) represent a more general class of abstract nouns, or *shell nouns* (Hunston & Francis, 2000; Schmid, 2000). The differences between the two groups suggest that, unlike the high-rated papers, the low-rated texts not only relied on a narrower range of nominalization types, but they also included the same words repeatedly.

Table 4 Top twenty most frequent nominalization types in the two corpora

	HC			LC		
Rank	Type	*n*	%	Type	*n*	%
1	student	2.71	7	student	5.97	15
2	government	1.66	4	information	2.34	6
3	information	1.37	3	government	.99	3
4	solution	1.08	3	reader	.91	2
5	situation	.79	2	activity	.81	2
6	population	.69	2	behaviour	.67	2
7	violence	.58	1	situation	.67	2
8	advertisement	.54	1	player	.56	1
9	education	.53	1	ability	.53	1
10	marriage	.48	1	health	.48	1
11	security*	.48	1	reporter	.48	1
12	activity	.45	1	attention	.46	1
13	organization	.42	1	difference	.46	1
14	development	.41	1	production	.46	1
15	ability	.39	1	education	.45	1
16	consumer*	.39	1	teacher	.43	1
17	evidence*	.39	1	definition*	.42	1
18	definition*	.38	1	truth	.40	1
19	leader	.38	1	evidence*	.39	1
20	action	.37	1	printer	.39	1
Total		14.49	36		18.27	47

Note: ***n*** = normed frequencies per 1,000 words in the corpus; % = percentage of the total instances of nominalization; * = nominalizations in the AWL (Coxhead, 2000).

To determine whether these top twenty nominalizations in both corpora are considered academic, we searched these words in the AWL (Coxhead, 2000). We found that four nominalization types (*security*, *consumer*, *evidence*, *definition*) in the HC appear on the AWL, accounting for 1.64 ptw total occurrences, in comparison to two (*evidence*, *definition*) in the LC, totalling.81 ptw instances. With the inclusion of a greater number of lexical items more appropriate for the academic prose register, the high-rated students

demonstrated, at least with the use of nominalization, their greater register awareness.

Abstract/concrete and human/non-human nominalization

As suggested from the differences in common suffixes used, the HC includes abstract nominalizations more frequently than the LC (see Table 5). It also contains fewer concrete nominalizations than the LC. Although the comparisons were not statistically significant, the frequency differences suggest that high-rated writers tend to place a greater focus on abstract concepts, thus reducing the emphasis on agents of action or concrete objects with physical representations.

Concrete nominalizations were further classified into human and non-human. Compared to the high group, the low group displayed a greater use of human nominalizations with the small effect size. This finding is supported by Table 4 above, in which five out of the top twenty nominalizations in the LC are human entities (*student*, *reader*, *player*, *reporter*, *teacher*). As Liardét (2016) indicates, 'academic discourse privileges impersonal language that moves away from naming or describing the actors' (p. 27). In our follow-up analysis of the more generically (and commonly) used words to express human agency (i.e., *person(s)*, *people*, *you*, *we*), we also found that the HC (*Mdn* = 9.29, *IQR* = 14.39) includes these four agency markers less frequently than the LC (*Mdn* = 14.26, *IQR* = 18.51). This shift from the naming and description of human agents toward abstract concepts suggests a developmental progression for nominalization. As students' writing proficiency increases, their reliance on human agents in their writing decreases while impersonalization increases.

Table 5 Descriptive statistics and results of Mann-Whitney U for abstract/concrete and human/non-human categories in the two corpora

	HC		LC			
Categories	*Mdn*	*IQR*	*Mdn*	*IQR*	*p*-value	*r*
Abstract	25.17	14.91	23.23	14.69	.184	.12
Concrete	9.88	10.02	11.96	19.16	.176	-.12
Human	6.45	7.97	8.42	18.64	.109	-.14
Non-human	1.43	4.99	1.46	3.34	.740	-.03

Note: Mdn = median normed frequencies per 1,000 words; *IQR* = interquartile range; statistical significance with Bonferroni correction is at $p<.013$.

Nominal stance

Overall comparisons of the stance options of nominalization show no statistically significant differences between the two groups for any categories,

as Table 6 shows. Even the proportional distributions of the categories do not show marked differences. The present study's findings indicate that both groups rely on nominalizations having similar stance types and that perhaps the stance of the nominalizations was not an aspect that affected the instructors' judgement of text quality. However, high-rated writers tended to use more nominalizations expressing event, cognition, and quality, although the effects were small. Greater reliance on the event stance in the HC aligns with previous studies on published writing (e.g., Jiang & Hyland, 2015), which demonstrate that this stance type is common in the academic register. Supporting previous research on L2 student writing (Lee & Deakin, 2016), this suggests that, by using this stance type more commonly, the high-rated students attempted to build their argument on factual evidence with the aim of persuading their audience. These writers, we argue, have a greater awareness that academic readers are likely to accept their arguments when the nominalizations take an event-oriented stance. The high-rated writers also relied slightly more on cognition (e.g., *solution*, *attention*, *decision*) and quality nouns (e.g., *truth*, *pressure*, *development*), which allow writers to discuss mental processes involved in events and to evaluate and show their positionality towards the evidence presented. This combination of stance types may have contributed positively to the assessment of their writing.

Table 6 Descriptive statistics and results of Mann-Whitney U for stance categories for abstract nominalizations in the two corpora

	HC		LC			
	Mdn (*IQR*)	%	*Mdn* (*IQR*)	%	*p*-value	*r*
Entity	17.52 (11.54)	74	15.18 (11.68)	74	.225	.11
Text	.00 (.00)	0	.00 (.00)	0	.445	.07
Event	12.30 (9.63)	54	10.93 (12.32)	50	.264	.10
Discourse	1.21 (2.29)	9	1.75 (3.37)	14	.734	.03
Cognition	2.08 (3.37)	11	1.64 (3.14)	10	.255	.10
Attribute	5.11 (4.86)	23	4.07 (3.05)	23	.121	.13
Quality	2.44 (3.51)	14	2.01 (1.94)	13	.058	.16
Manner	.92 (1.70)	5	.87 (1.88)	6	.920	.01
Status	.78 (1.50)	4	.82 (1.56)	4	.861	-.02
Relation	.17 (.94)	3	.42 (1.34)	3	.735	-.03
Total	25.17 (14.91)	100	23.23 (14.69)	100	.184	.12

Note: ***Mdn*** = median normed frequencies per 1,000 words; ***IQR*** = interquartile range; statistical significance with Bonferroni correction is at ***p***<.005.

Noun phrase modification

Among all classifications employed in this study, the comparison of noun modification types revealed the most notable differences between the two corpora. Table 7 shows that the HC contained more modified nominalizations than the LC with the small effect sizes. The effect size for the absence of modification was among the highest in our study, which suggests that this variable is likely to have contributed to the instructors' judgement of text quality. These findings align with research on syntactic complexity, which suggests that the frequency of modification in the noun phrases increases with writers' development or writing quality (e.g., Casal & Lee, 2019; Crossley & McNamara, 2014; Staples et al., 2016). Thus, the high-rated students' reliance on modified nominalizations indicates that the more grammatically complex texts they produced could have contributed to their higher grades.

When specific categories of modification were examined, more evidence of greater syntactic complexity in the HC was found. The high-rated writers used complex premodifications as well as combined pre- and postmodifications more frequently with small effect sizes. This indicates that high-rated papers often use more than one modifier with nominalizations, thus compressing more information into the noun phrase. Greater use of total premodification, total postmodification, attributive adjectives, and prepositional phrases, with small effect sizes, also shows that the HC includes more modified nominalizations, thus packaging more descriptive information into correspondingly complex noun phrases.

A difference with a small effect size was found in the use of possessive nouns as premodifiers, with the LC containing more of this feature. Considering that nouns in the possessive form are normally those referring to human beings, the greater use of possessives as premodifiers supports this study's previous findings that low-rated writers place a greater focus on human agents. This also suggests that high-rated writers possess a greater awareness that academic prose tends to focus more on abstract concepts rather than human agents. However, the findings related to possessive nouns, complement clauses, and appositives should be interpreted with caution since the frequencies of these features were rather low in both corpora.

Table 7 Descriptive statistics and results of Mann-Whitney U for modification categories in the two corpora

	HC		LC			
Categories	*Mdn*	*IQR*	*Mdn*	*IQR*	*p*-value	*r*
No modification	13.36	7.78	16.46	15.48	.069	-.16
Premodification	10.69	9.11	9.35	6.11	.257	.10
Attributive adjectives	5.27	6.97	4.55	4.66	.160	.12
Nouns	2.03	2.66	1.69	2.08	.504	.06
Possessive nouns	.00	.96	.83	1.11	.019	-.20
Complex premodification	1.09	2.44	.92	1.96	.225	.10
Postmodification	10.49	7.55	9.19	6.29	.234	.10
Prepositional phrases	5.01	4.94	4.58	3.61	.252	.10
Relative clauses	.80	1.51	.92	1.58	.679	-.04
Non-finite relative clauses	.00	.75	.00	.98	.635	-.04
Complement clauses	.00	.75	.42	1.02	.039	-.18
Appositives	.00	.00	.00	.00	.122	.13
Pre- and postmodification	3.21	3.39	2.55	2.54	.115	.14

Note: Mdn = median normed frequencies per 1,000 words; *IQR* = interquartile range; statistical significance with Bonferroni correction is at $p<.004$.

When we examined the top three most common nominalizations for the absence or presence of modification, the high-rated writers modified these words to a greater extent. *Student* was modified in 87% cases in the high-rated corpus as opposed to 46% in the low corpus; *information* was modified in 77% and 69% instances, respectively; and *government* in 63% and 50% cases, respectively. These findings demonstrate that although both groups relied on the same top three nominalization types, the high-rated writers tended to modify these nominalizations more frequently to make the noun phrases containing nominalizations more specific, informative, and complex, and thus reflective of academic writing:

1) Furthermore, according to Lazear (1995), the language environment would encourage international students, who belong to minority groups, to learn a second language more actively.
2) Therefore, readers still need to face much inaccurate information about gun control everyday.

3) Moreover, the central government of the People's Republic of China sent an army to stay in the islands which also became part of the Chinese Xisha citizens (Chen, 1987).

Conclusion

This study compared nominalization use in high- and low-rated L2 undergraduate research papers. The overall frequency of nominalizations was similar in both corpora when a comprehensive list of nineteen suffixes was used to identify the target features. However, slightly different results were obtained when a smaller set of five suffixes, commonly used in nominalization research, was examined. These differences demonstrate the importance of examining a full range of suffixes in order to obtain a more complete understanding of nominalization use. The more frequent use of nominalizations with five suffixes in the HC supports previous research (e.g., Grant & Ginther, 2000). Both groups relied on the full range of suffixes used to create nominalizations, and several suffixes (i.e., *-nt*, *-th*, and *-or*) that have not been typically examined in nominalization research were used proportionally more commonly than some that have been included.

No significant differences were found in any of the examined classifications, suggesting that high- and low-rated student writers are similar in their use of nominalization in terms of frequencies, unique types, abstract/concrete and human/non-human categories, nominal stance, and modification types. Previous studies on EAP writing have found differences across years of study, discipline, and genre (e.g., Staples et al., 2016), but considering the fact that the present study's writers were at the same educational level, taking the same writing course (though with different teachers and in different semesters), wrote the same essay type, and enrolled in the course based on similar entrance requirements, it may not be surprising that no statistical differences were found. However, the small effect sizes for certain categories point to subtle differences between the two corpora, which together with other aspects might have affected the instructors' evaluation of the content, organization, and academic language use of the students' texts. The HC was found to rely on a broader range of nominalization types. This finding was also supported when the twenty most frequent nominalization types were examined: while the low-rated writers tended to repeat the same nominalization types, the high-rated students demonstrated a wider repertoire of lexical choices. More nominalizations from the top twenty list also appeared on the AWL (Coxhead, 2000) in the HC, pointing to these writers' greater awareness of lexis expected in academic writing.

Greater awareness of academic conventions in the high-rated texts was also realized in several other ways. Unlike the low-rated students, the

high-rated writers relied more on abstract nominalizations as opposed to concrete human nominalizations. As underscored by other studies, this shift from human actors towards impersonal language is a characteristic feature of academic writing (Biber et al., 1999; Liardét, 2016). The HC contains the event stance noun type more frequently, which aligns with previous research on L2 student writing (Lee & Deakin, 2016) and published research articles (Jiang & Hyland, 2015), as well as cognition and quality types, showing mental processes and embedding evaluation of proposition. The high-rated writers modified the nominalizations more frequently overall and included more than one modifier, which demonstrates that these students package more information into noun phrases and present more detailed information about the nominalization. These findings support previous research that reveals greater use of noun phrase modification with increasing proficiency and experience (e.g., Casal & Lee, 2019; Crossley & McNamara; Staples et al., 2016). The overall results suggest that a greater frequency of modified abstract nominalization used to demarcate events (and to some extent cognition and quality) contribute to a positive assessment of L2 EAP student essays and may be potentially representative of higher academic writing proficiency.

While this study is among the few that examined nominalization using a finer-grained approach, going beyond general frequency counts, a few limitations exist. Although we attempted to mitigate the impact of different corpus sizes by normalizing the frequencies in each text, the two corpora are admittedly unbalanced. Thus, larger and more balanced sample sizes may contribute to more precise results in future research. Another variable that could have affected the results is the essay topics. While recognizing the difficulties of controlling variables including task, L1, and topic in naturalistic data, Caines and Buttery (2017) advise controlling for these variables 'as far as possible' (p. 22) to mitigate what they refer to as 'opportunity of use', or 'the opportunity the learner is afforded to use a linguistic feature' (p. 6). 'Where they cannot be fully controlled' (p. 22), they recommend acknowledging such limitations. Due to the nature of the corpus, topic was not controlled in study. The students wrote on a variety of topics, and in particular papers, irrespective of grade, some nominalization types were the keywords, which had to be repeated to maintain cohesion. Such texts contained much higher frequencies of nominalizations, and thus, minimizing the impact of topics should be considered in future research. Also, although most languages employ nominalization, variations exist in the way nominalizations are formed (Comrie & Thompson, 2007). The study's texts were produced by writers from a variety of linguistic backgrounds, but mostly L1 Mandarin speakers. Future research could make a greater effort to control for this variable to determine whether and to what degree the L1 impacts L2 English nominalization usage. Another

area that needs further analysis is the relationship between nominalization and textual cohesion in L2 student writing, as nominalization can serve as both anaphoric and cataphoric references to establish connections between ideas. Future research could tackle this issue using both quantitative and qualitative analyses to provide us with a more robust understanding of how L2 writers use nominalization to refer to propositions in the discourse to maintain cohesion.

This study has several pedagogical implications for L2 EAP writing instruction. While overall use of nominalizations was similar in the high- and low-rated essays, some differences were found between the two groups, which suggest that it may be important to devote time to instruction on nominalizations in L2 writing courses. Previous research has shown that students are often receptive to instruction on nominal features (e.g., Bychkovska, 2021; Casal & Lu, 2021) and that the skills they acquire during EAP instruction are perceived as transferrable and beneficial to their success in other courses they take (Zou & Jiang, 2021). EAP instructors may need to consider introducing students to the concept of nominalizations and their formation processes early in their education. Students should be made aware of the prominence of nominalization in contemporary academic writing (Biber & Gray, 2013), and of the important functions it serves, including making texts formal, objective, concise, and cohesive. Discussion of different types of nominalizations, such as abstract or concrete as well as the different types of nominal stance, can also enhance student writers' appropriate use of this linguistic feature. To raise students' awareness of how nominalizations perform various functions, teachers may engage students in noticing this feature in authentic published or high-rated student writing. Inevitably, the discussion of modification may also need to be introduced with the topic of nominalization, as noun phrase complexity is not only a distinctive linguistic characteristic of academic writing but also an important feature of successful L2 student writing (Casal & Lee, 2019). We argue that such instruction focusing on nominalization may lead novice EAP writers to gain better control over this important linguistic feature of academic writing and construct texts that academic readers would find cohesive and persuasive.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research received no specific grant for authorship and/or publication of this article.

References

Ädel, A. (2008). Involvement features in writing: Do time and interaction trump register awareness? In G. Gilquin, S. Papp, & M. B. Díez-Bedmar (Eds.), *Linking up contrastive and learner corpus research* (pp. 35–53). Rodopi.

Anthony, L. (2018). AntConc (version 3.4.7) [Computer Software]. Waseda University. http://www.laurenceanthony.net.

Banks, D. (2008). *The development of scientific writing: Linguistic features and historical context*. Equinox.

Baratta, A. M. (2010). Nominalization development across an undergraduate academic degree program. *Journal of Pragmatics*, *42*(4), 1017–1036.

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

Biber, D., & Gray, B. (2013). Nominalizing the verb phrase in academic science writing. In B. Aarts and J. Close (Eds.), *The verb phrase in English: Investigating recent language change with corpora* (pp. 99–132). Cambridge University Press.

Biber, D., & Gray, B. (2016). *Grammatical complexity in academic English*. Cambridge University Press.

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, *45*, 5–35.

Biber, D., Gray, B., & Poonpon, K. (2013). Pay attention to the phrasal structures: Going beyond T-units. A response to WeiWei Yang. *TESOL Quarterly*, *47*, 192–201.

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

Bychkovska, T. (2021). Effects of explicit instruction on noun phrase production in L2 undergraduate writing. *Journal of English for Academic Purposes*, *54*, 101040.

Caines, A., & Buttery, P. (2017). The effect of task and topic on opportunity of use in learner corpora. In V. Brezina & L. Flowerdew (Eds.), *Learner corpus research: New perspectives and applications* (pp. 5–27). Bloomsbury.

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

Casal, J. E., & Lu, X. (2021). "Maybe complicated is a better word": Second-language English graduate student responses to syntactic complexity in a genre-based academic writing course. *International Journal of English for Academic Purposes: Research and Practice*, *1*(1), 95–114.

Comrie, B., & Thompson, S. (2007). Lexical nominalization. In T. Shopen (Ed.), *Language typology and syntactic description* (2nd ed., pp. 334–381). Cambridge University Press.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*, 213–238.

Crossley, S., & McNamara, D. (2014). Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners. *Journal of Second Language Writing*, *26*, 66–79.

Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*, *22*, 166–178.

Derewianka, B. (2003). Grammatical metaphor in the transition to adolescence. In A.-M. Simon-Vandenbergen, M. Taverniers, & L. J. Ravelli (Eds.), *Grammatical metaphor: Views from systemic functional linguistics* (pp. 185–219). John Benjamins.

Gardner, S., Nesi, H., & Biber, D. (2019). Discipline, level, genre: Integrating situational perspectives in a new MD analysis of university student writing. *Applied Linguistics*, *40*, 646–674.
Grant, L., & Ginther, A. (2000). Using computer-tagged linguistic features to describe L2 writing differences. *Journal of Second Language Writing*, *9*(2), 123–145.
Halliday, M. A. K. (1989). *Spoken and written language*. Oxford University Press.
Hinkel, E. (1997). Indirectness in L1 and L2 academic writing. *Journal of Pragmatics*, *27*(3), 361–386.
Hunston, S., & Francis, G. (2000). *Pattern grammar*. John Benjamins.
Jiang, F., & Hyland, K. (2015). "The fact that": Stance nouns in disciplinary writing. *Discourse Studies*, *17*, 529–550.
Lancaster, Z. (2014). Exploring valued patterns of stance in upper-level student writing in the disciplines. *Written Communication*, *31*, 27–57.
Lee, J. J., & Deakin, L. (2016). Interactions in L1 and L2 undergraduate student writing: Interactional metadiscourse in successful and less-successful argumentative essays. *Journal of Second Language Writing*, *33*, 21–34.
Leedham, M. (2015). *Chinese students' writing in English: Implications from a corpus-driven study*. Routledge.
Liardét, C. L. (2013). An exploration of Chinese EFL learner's deployment of grammatical metaphor: Learning to make academically valued meanings. *Journal of Second Language Writing*, *22*, 161–178.
Liardét, C. L. (2016). Nominalization and grammatical metaphor: Elaborating the theory. *English for Specific Purposes*, *44*, 16–29.
Norman, G. J. (2003). Consistent naming in scientific writing: Sound advice or shibboleth? *English for Specific Purposes*, *22*, 113–130.
Ryshina-Pankova, M. (2010). Toward mastering the discourses of reasoning: Use of grammatical metaphor at advanced levels of foreign language acquisition. *The Modern Language Journal*, *94*, 181–197.
Schmid, H. J. (2000). *English abstract nouns as conceptual shells: From corpus to cognition*. Walter de Gruyter.
Staples, S., Egbert, J., Biber, D., & Gray B. (2016). Academic writing development at the university level: Phrasal and clausal complexity across level of study, discipline, and genre. *Written Communication*, *33*, 149–183.
Thompson, G. (2010). Grammatical metaphor and success in academic writing. In S. Hunston & D. Oakey (Eds.), *Introducing applied linguistics: Concepts and skills* (pp. 27–33). Routledge.
Yoon, C. (2018). Nominalization in Korean EFL learners' argumentative writing: A comparative study of distribution and use. *The New Studies of English Language & Literature*, *69*, 249–274.
Zou, B., & Jiang, G. (2021). The impact of EAP skills on students' academic study. *International Journal of English for Academic Purposes: Research and Practice*, *1*(1), 57–80.

Biographies

Tetyana (Tanya) Bychkovska, M.A., is an independent scholar, whose research interests include second language academic writing, corpus linguistics, and lexico-grammar. Her work has appeared in *Journal of English for Academic Purposes* and *System*.

Joseph J. Lee, PhD, is a senior lecturer in English linguistics in the Department of English at Dalarna University. His research and teaching interests include ESP/EAP, academic writing, genre studies, classroom discourse, applied corpus linguistics, and teacher education. His publications have appeared in *English for Specific Purposes*, *Journal of Response to Writing*, *Journal of Second Language Writing*, and *System*.

Appendix
List of derivational suffixes analysed

Suffix	Examples
-acy	*supremacy*, *legitimacy*
-age	*storage*, *usage*
-al	*arrival*, *survival*
-dom	*freedom*, *wisdom*
-ee	*employee*, *interviewee*
-er	*abuser*, *driver*
-ion	*decision*, *education*
-ism	*criticism*
-ist	*activist*, *feminist*
-ment	*abandonment*, *empowerment*
-nce	*independence*, *maintenance*
-ncy	*dependency*, *expectancy*
-ness	*awareness*, *consciousness*
-nt	*consultant*, *entrant*
-or	*actor*, *director*
-th	*growth*, *strength*
-ty	*anxiety*, *safety*
-ure	*conjecture*, *exposure*
-y	*jealousy*, *victory*