

# Triangulated Corpus-Informed Assessment of Multilingual Academic Writing

Intan Pradita\*<sup>1</sup>, Adi Sutrisno<sup>2</sup>, Tofan Dwi Hardjanto<sup>3</sup>, Cihad Gündüz<sup>4</sup>

<sup>1</sup>Universitas Islam Indonesia, Yogyakarta, Indonesia

<sup>2,3</sup>Universitas Gadjah Mada, Yogyakarta, Indonesia

<sup>4</sup>Dicle University, Diyarbakir, Turkey

## ABSTRACT

The naturalness of formulaic language is a key variable in the acceptability of multilingual academic writing. Such naturalness is best reflected by lexical collocations between linguistic units. However, no well-established framework has systematically studied collocational acceptability in multilingual academic contexts. This study fills the gap by proposing a triangulated model to assess academic lexical collocations produced by Indonesian authors in English Language Teaching (ELT) research publications.

The model integrates three components: corpus data, statistical metrics, and cross-group speaker judgments based on Kachru's (1985) three circles of English. Findings show that many recurrent collocations, such as specific Adverb-Verb and Verb-Noun structures, are semantically transparent and moderately acceptable, even when they diverge from prototypical native-speaker norms. By applying a decision logic of MI Score > 3.0, a negative T-score, and I-CVI > 0.78, the proposed model contributes to research on corpus linguistics in education by linking phraseology competence and multilingual academic communication.

## ARTICLE INFO

*Paper type:*  
Research Article

*Article history:*  
Received: 13/04/2026  
Revised: 25/05/2026  
Accepted: 02/06/2026

*Keywords:*

- academic writing
- corpus-informed language assessment
- lexical collocations
- semantics
- phraseological competence

## 1. Introduction

The number of non-Anglophone English speakers are more than those of Anglophone speakers. This way brings English as a lingua franca across both spoken and written discourse. Within academic context, Mauranen (2020) postulates an approach of English as an Academic Lingua Franca. The approach argues that English is functioned as the medium through which worldwide scholars deliver their research findings and logic-bound argumentation. In the Global South context, this phenomenon creates a significant linguistic hurdle. Indonesian scholars, for example, face a challenge in publishing research findings in English-medium journals. Whereas Indonesian scholars represent one such growing demographic, striving to contribute to international knowledge production.

The challenge centers on acceptability, is closely tied to intelligibility, a multifaceted concept that transcends grammatical accuracy (Flowerdew, 2008; Mauranen, 2012) and encompasses the regularities that make academic writing recognizable and natural to its readership (Mauranen, 2020). The mechanism that is recognized to achieve the naturalness is lexical collocations. It is the conventional co-occurrence of words that signals fluency to expert readers (Wray, 2013). Since Indonesian authors are non-Anglophone writers, unconventional collocations may rise not from grammatical failure but from unconventional or unfamiliar word combinations. Thus, although a word combination is syntactically correct, it feels opaque to peer reviewers. This friction stems from what Aybek (2025), Demir (2017), and Mauranen (2012) identify as L1 interference, lexical unfamiliarity, and cultural influences such as communality (Arsyad & Adila, 2018) and politeness norms (Triyoko et al., 2021).

Corpus linguistics provides the richest empirical basis for testing collocational naturalness. First, Götz (2022) proves that learner corpora are ecologically valid, strike a balance between intuition and empirical evidence, and detect linguistic markers that correlate with certain levels of proficiency. Methodologically, there are two key elements consistently observed in collocational research in non-native academic texts: the former uses Mutual Information (MI) scores to calculate the statistical strength of collocation, that is, the extent to which two words co-occur more often than would be expected by chance (Mauranen, 2020; Gablasova et al., 2017), and the latter applies t-tests to test the difference between non-native and native corpus statistics for significance (Aybek, 2025; Demir, 2018). Following this protocol, Zhao (2023) has tested collocations in English abstracts by Chinese authors against native-authored English; Yilmaz and Römer (2020) have done the same with grammatical collocations. Any statistically significant deviations identified through such testing indicate collocational patterns in question as non-prototypical and therefore worthy of editing. To this extent, corpus-informed assessments are proven to accurately rate text readability, as shown by Crossley et al. (2023, 2024); the feasibility of the approach in the domain of spoken discourse is confirmed by Taguchi (2009).

<sup>1</sup> [intan.pradita@uii.ac.id](mailto:intan.pradita@uii.ac.id) (Pradita)

There is, however, a limitation of corpus studies. It is that statistical deviation from native-speaker norms does not constitute unacceptability. A collocation that is less frequently used by native English speakers but remains semantically intelligible becomes the focus of this study. The corpus-statistical approach identifies unfamiliar collocations but it cannot deliver an acceptability verdict. This is due to corpus linguistics has limited equipment on statistical measurement of acceptability. However, corpus linguistic tools such as SketchEngine provides a statistical measurement of collocation strength to show how consistent the word combination is. Whereas there are two urgencies to address. First, the needs of additional aids to assess English lexical collocations that embody the norms of English as a Lingua Franca instead of the strong persistence on native-speakerism norms. By then, the assessment perceives unfamiliar collocations as the impact of English used across a globally diverse community (Kachru et al., 2006; Götz, 2022). Second, the needs to meet the level of tolerance in the acceptability rate of lexical collocations among English speakers (Pradita & Puspitasari, 2025). By doing so, the concept of acceptability is reliable. In this respect, there is gap point to a need for a framework that is simultaneously corpus-grounded, statistically rigorous, and sociolinguistically inclusive.

In this respect, it is human judgment which offers the interpretive element that corpus-based approaches fail to offer on their own. This point has been made clear through studies conducted by Crossley et al. (2023; 2024) and Götz (2022). They suggest the importance for raters to contribute to evaluations of texts that go beyond what can be achieved by any computer program. The involvement of judgment on behalf of English speakers with different backgrounds is crucial, because acceptability works within real-world ELF scenarios. Thus, this research will rely on the judgment of English-speaking raters belonging to all of Kachru's (1985) three circles of English.

This study aims at filling the void by presenting a pilot triangulated approach that combines three criteria for collocational acceptability. The three aspects are MI scores for the statistical component, t-test results for the comparative component, and judgments made by English-speaking individuals from diverse linguistic backgrounds for the evaluative component. By combining the three factors, the collocation can be classified as acceptable. This triangulation model is statistically significant, not deviating from the normative patterns observed in the corpus and rated as natural by the reader sample that represents diverse linguistic communities. The convergence of these criteria constitutes the conceptual contribution of the present framework and is defined here for the first time as the definition of collocational acceptability in ELF contexts. The current study aims to address the following research questions:

1. How does the triangulation model combine corpus linguistics, statistic measures, and human judgement to assess Phraseological competence in non-native English academic writing?
2. How does the triangulation model define acceptability criteria of a lexical collocation?

It should be noted that, as a pilot model, this study focuses on subject-specific collocations within the scope of English Language Teaching. This disciplinary narrowing ensures comparability of rhetorical conventions across the two writing communities, though findings should be interpreted within this boundary. English Language Teaching was selected because it offers a well-defined and relatively homogeneous academic genre, marked by consistent rhetorical structures and shared discourse conventions.

## 2. Methods

### 2.1. Data

This study utilized two primary datasets to depict the distinct writing groups. The main data was the PRAYA Corpus (Pradita & Puspitasari, 2025) which consists of a collection of Indonesian-authored research publications in English. There were 32 Scopus-indexed publications related to English Language Teaching. The corpus has 112,322 tokens. As the corpus reference, this study utilized the PRANA Corpus. It is a dataset developed by the authors consisting of 31 research articles written by Anglo American authors. It consists of 139,206 tokens. This study utilized two primary datasets to depict the distinct writing groups. The main data was the PRAYA Corpus (Pradita & Puspitasari, 2025) which consists of a collection of Indonesian-authored research publications in English. There were 32 Scopus-indexed publications related to English Language Teaching. The corpus has 112,322 tokens. The reference corpus, PRANA, was developed by the authors consisting of 31 research articles written by Anglo American authors. It consists of 139,206 tokens. The two corpora are comparable in terms of genre, disciplinary focus, and size. Although the token ratio of the PRAYA and PRANA is 1: 1,24, it falls within the acceptable range of comparable corpus (McEnery & Hardie, 2012). Below is the inclusion and exclusion criteria of the PRAYA and PRANA.

Table 1. The Inclusion and Exclusion Criteria

Type	Criterion	Rationale
Inclusion	Research articles in ELT that are published in reputable international journals with impact factor > 0.25  The articles were published in 2018-2023	Aligned with the Ministry of Education, Research, and Technology (2019) policy requiring scholars with 10–20 years of experience to publish in reputable journals. Experienced authorship ensures more consistent phraseological patterning.  The time boundary of 2018-2023 partly to decrease the probability that the articles we collected were written with the help of AI. Large language models and other generative AI tools capable of producing academic writing were not widely available until late 2022.
Exclusion	ELT articles published in proceedings or predatory journals. PRAYA: Articles co-authored with international (non-Indonesian) scholars PRANA: Articles co-authored with Indonesian scholars	Possible insufficient peer review.  Collaboration with native or international speakers may introduce external linguistic influence, compromising the integrity of the Indonesian-authored writing sample.

## 2.2. Tools

The analytical tool used in this research was Sketch Engine. It was used to extract lexical collocations from both corpora and to calculate Mutual Information (MI) Scores for each collocation. Only collocations with MI score > 3.0 were kept for further analysis, following the conventional threshold for statistically meaningful co-occurrence (Hunston, 2002).

## 2.3. Procedures

### 2.3.1. Identifying Lexical Collocations

Lexical collocations were identified across four grammatical categories: Adjective-Noun, Adverb-Verb, Adverb-Adjective, and Verb-Noun. Collocations from the PRAYA Corpus were extracted using Sketch Engine and filtered by MI Score > 3.0.

### 2.3.2. Corpus Comparative Analysis

The retained collocations were compared against the PRANA Corpus using an independent samples T-test to establish whether PRAYA collocations deviated significantly from Anglo-American writing norms.

### 2.3.3. Measuring Expert Validity

To make sure that the lexical collocations have been correctly categorized and globally acceptable, this study did expert validity to four speakers of English. Those are a NS from England, an EFL NNS from Turkey, an EFL NNS from Indonesia, and an ESL NNS from Philippine. The experts assessed the intelligibility of the ten samples of lexical collocations from each type of collocations as found in the PRAYA corpus. The researcher used three Likert scales of acceptability rate ranging from acceptable to unacceptable. Fourth, the researcher analysed the ratings quantitatively by doing inter-rater reliability test. The score from each expert was calculated based on the Item-level Content Validity Index (I-CVI). We calculated the scores by dividing the number of experts giving ratings of 0-2 by the total number of experts (Polit & Beck, 2006). If the score is more than 0,78, the collocation is acceptable. If the score is below 0,70, the collocation should be reconsidered.

## 2.4. Analysis

To reach this stage, the acceptable collocations were interpreted by considering the discourse of social interaction (Hyland & Tse, 2004), and current patterns of Indonesian-authored English writings (Arsyad & Adila, 2018). The interpretation was conducted to determine the extent accommodation in ELF Writing. Collocational acceptability was decided based on a three-way decision logic reflecting the three components of the triangulation model: corpus salience, statistical comparison, and human judgement. A collocation was deemed acceptable when it met all the following criteria. Those are MI Score > 3.0, negative T-score with respect to the PRANA Corpus, and I-CVI of  $\geq 0.78$ . The I-CVI was calculated by dividing the number of experts who gave acceptable ratings by the total number of raters (Polit & Beck, 2006).

I-CVI was selected over other inter-rater measures because it assesses acceptability at the item level, reflecting the percentage of agreement among experts for each collocation individually. The I-CVI, unlike Cohen's Kappa or Fleiss' Kappa, does not assume rater interchangeability, making it particularly appropriate for a cross-circle panel in which each rater's linguistic background is a different structurally distinct variety of English. An I-CVI cut-off of  $\geq 0.78$  was taken as a marker of sufficient cross-group consensus for a collocation to be considered acceptable in ELF contexts (Polit & Beck, 2006). Collocations with scores below 0.70 were flagged for re-consideration. This measure was preferred to simple inter-rater agreement because it reflects item-level consensus across raters from different linguistic backgrounds. This distinction is particularly relevant in ELF contexts where acceptability varies across English-speaking communities.

### 3. Results and Discussion

#### 3.1. The Triangulation Model of Acceptable Lexical Collocations

Among the four categories of lexical collocations, Table 1 summarizes the findings of each collocation that have similar nodes. Table 1 presents the distribution and strength of the three types of lexical collocations identified within the corpus: adjective-noun, adverb-verb, and adverb-adjective word combinations. The information provided includes the number of word combinations for each type, examples from the corpus, and the average Mutual Information (MI) score for each type, which represents the strength of the word combination.

Table 2. The lexical collocations in the PRAYA Corpus

Type of Collocations	Adjective-Noun	Adverb-Verb	Adverb-Adjective
The Sample	..and guiding them into an <i>English capability</i>	The second question was dealing with the kinds of social media that are <i>frequently used</i> by the students.	... after learning activities were conducted using Moodle workshop activity module were <i>significantly different</i> .
MI-Score	5,15	7,73	7,74
	After the module was valid one to one trial was done to three students with low, medium, and <b>high ability</b>	In English Language Teaching evaluation and assessment are among the most <b>frequently investigated</b> areas besides curriculum development	may be the only place where students learn English as it is a foreign language in Indonesia As Indonesian parents do not <b>commonly speak</b> English at home.
MI-Score	4,78	7,56	8,76

Based on the PRAYA Corpus, three types of collocation are identified. They are the Adjective-Noun, Adverb-Verb and the Adverb-Adjective types. The most common were Adjective-Noun combinations, which is a general tendency in academic writing to make use of descriptive modification for conceptual precision (Biber et al., 1999). Crucially, the average MI scores for all three types were higher than 3.0, suggesting moderate to strong collocational strength in the corpus. This conforms to Hunston's (2002) threshold for statistically significant co-occurrence and indicates that ELT scholars in Indonesia generate collocations that are not random pairings but exhibit steady phraseological patterns. In the lens of lexicogrammar, these types of collocations are sites of mutual constraining between lexical preference and grammatical structure. In academic writing, choices of words are not grammatically neutral because selecting a particular adjective or adverb at once triggers a set of grammatical expectations about the possible words to combine. The prevalence of Adjective-Noun structures in the PRAYA Corpus shows that Indonesian ELT scholars have internalized the nominal orientation of academic register.

The collocation of *English capability*, with an MI score of 5.15, illustrates a pattern where the Adjective-Noun structure is semantically transparent yet deviates from more prototypical native-speaker collocations such as *English proficiency* or *English competence*. Lexicogrammatically, *capability* belongs to a different collocational frame than *proficiency* or *competence*. It typically co-occurs with modal constructions expressing potential (the capability to do) rather than with language-descriptor adjectives such as English. Thus, the collocation of *English capability* in the lens of lexicogrammar tends to mismatch in which the grammatical slot is correctly filled but the lexical item selected does not fully conform to the register expectations of that slot. This kind of near-equivalent substitution has been documented in ELF writing as a form of productive accommodation rather than error (Mauranen, 2012; Seidlhofer, 2011). In the case of *high ability*, it shows a grammatically well-formed but lexically constrained Adjective-Noun

structure. In Anglophone academic writing, *ability* is more often found with prepositional phrases such as *ability to perform* or *ability in reading*. It also appears with intensifiers that indicate degree rather than scale such as *strong ability* or *limited ability*. The use of *high* as a scalar modifier is a lexicogrammatical pattern more appropriate for measurement contexts such as *high score* than for description of human capacity. However, the acceptable I-CVI ratings suggest that the grammatical transparency of the Adjective-Noun frame is sufficient for cross-circle intelligibility. To this extent, even if the particular lexical pairing is unconventional, it supports Mauranen's (2012) argument that intelligibility rather than native-speaker conformity governs acceptability in international academic discourse.

The collocation *frequently investigated* involves a more complex lexicogrammatical pattern. Although *frequently* as an adverbial modifier is frequent in academic texts, *investigated* as a verb head is not typical in such a collocation, in comparison with *widely investigated* or *extensively studied*, for example. From the lexicogrammatical point of view, *frequently* prefers to combine with verbs indicating stative or habitual processes such as *frequently occurs* or *frequently reported*. While *investigate* is a material process, which should be modified by scope adverbs such as *widely* and *extensively*. In other words, there is a discrepancy between the grammatical meaning potential of an adverb and the actual use. To this extent, there is a characteristic of the incomplete acquisition of constraints of adverb-verb collocations in English academic writing. In the Adverb-Adjective collocation, *commonly speak* can be considered as an example of lexicogrammatical transfer from Indonesian rhetorical structures. In English academic writing, the word *commonly* grammatically collocates to stative and attributive constructions such as *commonly used*, *commonly found*, or *commonly associated*. Whereas the word *speak* is a dynamic verb and thus indicates an active material process. Within the findings, these mismatches likely reflect an L1 transfer pattern in which Indonesian frequency adverbials apply more broadly across verb process types without the register-specific constraints that govern their English equivalents (Arsyad & Adila, 2018). Its lower I-CVI ratings confirm that cross-circle raters perceived this lexicogrammatical incongruence, reinforcing the triangulation model's capacity to identify collocations that fall outside the accommodative range of ELF academic writing.

### 3.1.1. The MI-score

The first aspect of the triangulation method is to seek the MI-score of the lexical collocations provided by the SketchEngine. Table 2 displays the collocations, the concordances, and the MI-score.

Table 3. The MI-Score of *English capability* and *frequently used* in the PRAYA Corpus

	PRAYA	
	Sample	MI -Score
Adverb-Adjective	The second question was dealing with the kinds of social media that are <i>frequently used</i> by the students.	7,73
Adjective-Noun	..and guiding them into an <i>English capability</i>	5,15

Referring to the criteria of collocational strength, Lei & Liu (2018) suggested that the MI-score should be not less than 3,00. In this case, if the MI-score stands by its own, *English capability* and *frequently used* are acceptable for it has a high collocational strength. We have not found whether the collocations are acceptable or natural in the context of English academic writing. Thus, it is necessary to move to the next aspect of the triangulation method that is corpus-comparative test by using T-test.

### 3.1.2. Corpus Comparative Analysis

The second aspect is corpus-comparative analysis. In this stage, we compared not only the mean of MI-score of the whole Adjective-Noun and Adverb-Verb collocations, but also the availability of the collocations. The results obtained through the analysis of the corpus, which compared the Indonesian-authored text, PRAYA, with the native-authored text, PRANA, show a marked trend of linguistic convergence in the academic ELT genre. To confirm this case, we did T-test for the mean of MI-score.

Table 3. T-test of the Lexical Collocations in the PRAYA and the PRANA Corpus

t-test for Equality of Means	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% CI Lower	95% CI Upper
Equal variances assumed	-1.18	11	0.26	-0.79	0.68	-2.29	0.71
Equal variances not assumed	-1.20	10.82	0.26	-0.79	0.66	-2.26	0.68

Independent samples t-test was used to test the difference in mean MI scores between two groups as shown in Table 3. Under the assumption of equal variances,  $t(11) = -1.18$ ,  $p = 0.26$ . The test was repeated without the assumption of equal variances and the result was very similar:  $t(10.82) = -1.20$ ,  $p = 0.26$ . The p-value was much larger than the conventional cut-off points of 0.05 in both cases, indicating that the difference in the mean MI scores was not statistically significant. The mean difference was -0.79, indicating that the average MI score for the first group was slightly lower than that for the second group. However, the 95% CI for the observed difference was from -2.29 to 0.71, and from -2.26 to 0.68, which included 0, suggesting that the observed difference was simply due to chance.

The non-significant findings of the t-test are in line with previous studies on collocational variation and phraseological norms, which have indicated that variations in the strength of association between words may follow gradients of conventionality rather than clear differences in usage (Aybek, 2025). This means that the non-significant findings of the t-test on MI scores suggest that the two sets of collocations do not differ significantly in terms of the strength of association, which may support the notion that such collocations may vary across corpora and may not necessarily point to any kind of deviance or aberrance in usage. From the point of view of ELF, the non-significant findings may suggest that multilingual writers of academic discourse can create phraseological patterns which are quantifiable but within the normal range of variation in expert writing (Mauranen, 2012). This suggests that the collocations produced by Indonesian learners were not significantly different from those of native speakers. However, the mean difference of -0.79 indicates that learners consistently scored lower than native speakers. To this extent, specific collocational deviations may still exist. To find out the possible deviations, we compared the availability of both collocations in the PRANA corpus. Table 4 shows the comparison of availability and the concordance.

Table 4. Results of the Corpus Comparison

	PRAYA	MI - Score	PRANA	MI -Score
Adverb- Adjective	...the most <b>commonly used</b> Vygotsky's learning concept in exploring parent-child interactions is the scaffolding theory	5,30	.. but other <b>commonly used</b> analyses (Rasch, MTMM, and factor analysis), are not right for the job.	7,33
Adjective- Noun	..and guiding them into an <b>English capability</b>	5,15	The primary purpose of this survey-based study was to measure the <b>English proficiency</b>	7,48

Table 4 informs us that in the case of availability, *English capability* is entirely absent in the PRANA corpus. Whereas, in the case of *commonly used*, the collocation is available in the PRANA corpus. The occurrence of these collocations is governed by both the established conventions of academic writing and their semantic appropriateness. It indicates that *English capability* is neither favorable for native English-speaking authors in ELT academic writing, nor semantically acceptable. The dataset in the PRANA corpus favors *English proficiency* more than *English capability*. Whereas, *commonly used* is either favorable or semantically appropriate in ELT academic writing. To this extent, favors is assumed to be an important aspect to define acceptability of lexical collocations. Thus, we did the third stage of triangulation model. It is inviting speakers from three circles of English to judge the acceptability of the lexical collocations.

### 3.1.3 Expert Validity

To continue the discussion, referring to Muraenen (2018), human validity is necessary to capture the acceptability based on English background. Table 4 shows the scoring from four speakers of three circles of English.

Table 5. Expert Judgement Result

No	Collocations	NS	NNS (ESL)	NNS (Ina)	NNS (EFL)	NoAR	I-CVI	Decision
1	Commonly used	2	2	2	0	4	0,75	Acceptable
2	English capability	1	0	0	0	4	0,25	Unacceptable

Table 5 shows the expert validations of twelve lexical collocations by using I-CVI scores. The raters represented Native Speakers as the main reference of Standard English, Non-Native Speakers of ESL users as the reference of outer circle English, and Non-Native Speakers of EFL users as the reference of expanding circle English. Eleven collocations are rated as Acceptable proven by the I-CVI of 0,75. However, there is one collocation of commonly given that has unacceptable rate proven by the I-CVI score of 0,25 (Polit & Beck, 2006). In the case of *English capability*, the word *capable* is unnatural to modify a language name like English. The institutionalized collocations to describe or modify someone's language ability are *proficiency*, *competence*, *skills*, or *ability*. Semantically, the word *capability* tends to modify a potential capacity of a system or an organization. The examples are military capability or cognitive capability. Whereas, in the case of *commonly used*, the word *commonly* is a frequent modifier of *used* in academic writing. As informed in the concordance, *commonly used* is functioned to emphasize the possibility of mushroomed use of social media (see Table 4).

### 3.1.4. Generating Model of Acceptability in ELF Writing

Referring to the previous findings and discussions, under the lens of Computational Linguistics, the idea of acceptability in ELF writing revolves around the strong Mutual Information score (MI Score), non-significant T-score interdiscursive hybridity (See Mauranen, et.al (2020)). It is an approach that accommodates cultural and linguistic diversity as an inseparable factor in ELF writing. The factors can be the use of indirect introductions to elaborate research background and hedging to show politeness (Mauranen, et al, 2020). In this research, Indonesian scholars tend to use adverbs of manner to emphasize actions (Adverb-Verb) and to strengthen remarkable findings (Adverb-Adjective). The pattern of emphasizing and strengthening actions and findings is included as boosters. Boosters in academic writing is to emphasize force of writers' certainty (Hyland & Tse, 2004). The use of boosters by ELF writers have been found in previous studies (Kobayashi, 2016 and Triyoko, et al, 2021). Although boosters are categorized as metadiscourse markers but in this study some the boosters are also considerable in the form of lexical collocations. To depict this concept, Figure 1 is proposed as the model of intelligibility and acceptability in ELF writing generated from the computational linguistics.

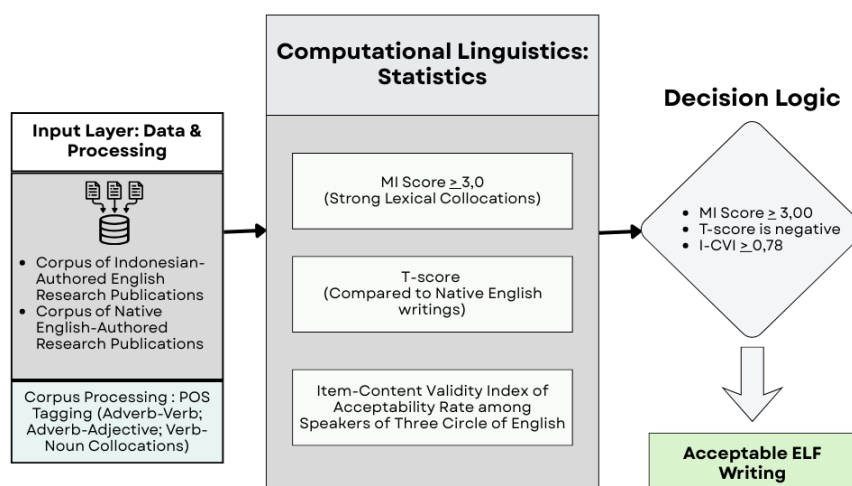


Figure 1. Acceptability of ELF Writing in the lens of Computational Linguistics

Figure 1 is the illustration of the whole process enacted through the data analysis. Collocations as one of linguistic unit is treated as a significant factor to determine the acceptability of ELF Writing. The

decision logic is when the collocations meet all thresholds such as strong MI score, negative T-score, and high I-CVI score. The model has some pedagogical implications, especially in English Language Teaching. First, the model as a combination of corpus statistics and cross-group acceptability judgements. The model will help English teachers in ELF countries to shift from rigid native-speakerism to acceptable phraseology. In compared to qualitative studies on ELF writing (see Huh, et al., 2020) have made a compelling argument for treating this kind of hybridity as emerging and legitimate rather than inadequate. Nevertheless, qualitative approaches to ELF writing do not provide replicable and quantitative criteria for assessing the legitimacy of phrases among non-native speakers. This theoretical framework provides, three measurable indicators of phraseological acceptability, which are MI score, T-score, and I-CVI. As an implication, this pilot model can help English teachers to activate evidence more than intuition.

### 3.2. Acceptability Criteria of Non-native English Phraseological Patterns

The triangulated model shows that the boundary between the natural and unnatural lexical choices or between the acceptable and unacceptable collocations is no longer tied only to grammatical correctness but also to collocational acceptability. Table 6 summarizes the criteria of non-native English phraseological patterns.

Table 6. Expert Judgement Result

No	Evidence	Weight
1	Collocation available in native English corpus	Strong indication of acceptable collocations
2	MI-score is more than 3,00	Suggest strong association
3	T-test is negative	No systematic difference, hence, the collocation is acceptable
4	I-CVI is more than 0,78	Strongly acceptable
5	I-CVI is more than 0,70 until 0,77	Almost acceptable

Referring to the findings, L1 rhetorical practice is complemented by international expectations of academic writing. The appraisal of English used by multilingual, or non-native, researchers is typically geared toward ensuring their specialist contributions are easy to grasp and compelling (Suhandoko, et al, 2025). To this extent, the area of intelligibility and acceptability in ELF Writing revolves solely around the phraseological and discourse functions. ELF Writing that contain semantically transparent lexical collocations and overuse boosters are still acceptable. Whereas the not acceptable one is those of writings that have incorrect grammar, weak lexical collocational strengths, and negative semantic transfer.

Further operationalization of the acceptability of lexical collocations in academic writing in ELF could be done via certain criteria as outlined in Table 6. The first and arguably the most basic criterion involves the presence of a particular collocation in a native English corpus. As pointed out by Nesselhauf (2005) and Sinclair (1991), this is an indicator that the phraseology in question follows conventions of conventionalized phraseology. Hence, when there is an example of usage of a particular collocation in large corpora of native speakers, including the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA), it should be considered part of standard usage patterns recognized by an international academic audience. However, the choice of using I-CVI as a tool for validating the acceptability of collocations in the current study is justified by the fact that it enables the integration of expert human judgment with the objective data obtained through corpus-based measures. The significance of this integration lies in the point that Xi (2017) additionally makes the case for important shortcomings associated with the gap between the capabilities offered by corpus linguistics and the ways in which language testing scholars apply them. Thus, in practice, corpora are used in validating tests based on benchmarks associated with native speaker and L2 learner populations. As such, this point highlights the role that should be attributed to the I-CVI criterion in the current study. While the former offers the criterion of benchmarked attestation in accordance with the normative usage standards (Mauranen, 2020; Suhandoko, et al, 2025), the I-CVI criterion performs an important complementary role. It allows English users to apply content validation as a criterion of expert human rater consensus. Thus, allowing one to validate collocations not only based on statistical but also pedagogically and communicatively relevant criteria

#### 4. Conclusion

This study found that there are traces of the influence of the Indonesian rhetoric on the choice of lexical items used to formulate academic claims. The influence is found that some collocations cannot be attributed to the English academic register. Regardless the interference, the lexical collocations are semantically transparent and mutually understandable. Even though most collocations were consistent with patterns typical of the PRANA corpus, it is necessary to consider the limitations in the findings. Firstly, the examined corpus includes only 32 articles related to one particular discipline. In addition, the four-expert rater panel is a relatively small sample of English speakers worldwide, which means that the results have not yet to be generalized as English academic writing. However, the triangulation approach proves that the degree of collocational acceptability in ELF academic prose varies on a continuum and encompasses completely acceptable collocations like "*significantly different*," near equivalent collocations like "*English capability*," and lexicogramatically inconsistent collocations like "*commonly speak*."

#### References

- Arsyad, S., & Adila, D. (2018). Using local style when writing in English: the citing behaviour of Indonesian authors in English research article introductions. *Asian Englishes*, 20(2), 170-185.
- Aybek, S. (2025). To Collocate or not to Collocate: Exploring Verb-Noun Collocations of Turkish EFL Learners. *Sage Open*, 15(4), 21582440251395720.
- Crossley, S. A., Tian, Y., Baffour, P., Franklin, A., Benner, M., & Boser, U. (2024). A large-scale corpus for assessing written argumentation: PERSUADE 2.0. *Assessing Writing*, 61, 100865.
- Crossley, S., Heintz, A., Choi, J. S., Batchelor, J., Karimi, M., & Malatinszky, A. (2023). A large-scaled corpus for assessing text readability. *Behavior Research Methods*, 55(2), 491-507.
- Demir, C. (2017). Lexical collocations in English: A comparative study of native and non-native scholars of English. *Journal of Language and Linguistic Studies*, 13(1), 75–87.
- Demir, C. (2018). Statistical measures in collocational studies: A corpus-based approach. *Applied Linguistics Review*, 9(1), 75–96.
- Flowerdew, J. (2008). Scholarly writers who use English as an Additional Language: What can Goffman's "Stigma" tell us?. *Journal of English for academic purposes*, 7(2), 77-86.
- Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. *Language Learning*, 67(S1), 155–179. <https://doi.org/10.1111/lang.12225>
- Goffman, E. (1968). *Stigma: Notes on the management of spoiled identity*. Penguin Books.
- Götz, S. (2022). Learner corpora to inform testing and assessment. In *The Routledge handbook of corpora and English language teaching and learning* (pp. 311-326). Routledge.
- Huh, M. H., Lee, I., & Kim, J. (2020). Written Academic ELF: Developing Writing Concepts in the New Normal. *English Teaching*, 75(4), 57-79.
- Hyland, K., & Tse, P. (2004). Metadiscourse in academic writing: A reappraisal. *Applied linguistics*, 25(2), 156-177.
- Kachru, Y. (2019). Speaking and writing in world Englishes. *The handbook of world Englishes*, 349-367.
- Kachru, B. B., Kachru, Y., & Nelson, C. L. (Eds.). (2006). *The handbook of world Englishes*. Blackwell.
- Kachru, B. B. (1985). Standards, codification, and sociolinguistic realism: The English language in the outer circle. In R. Quirk & H. G. Widdowson (Eds.), *English in the world: Teaching and learning the language and literatures* (pp. 11–30). Cambridge University Press.
- Kobayashi, Y. (2016). Investigating metadiscourse markers in Asian Englishes: A corpus-based approach. *Language in Focus: International Journal of Studies in Applied Linguistics and ELT*, 2(1), 19-35.
- Lei, L., & Liu, D. (2018). Research trends in applied linguistics from 2005 to 2016: A bibliometric analysis and its implications. *Applied Linguistics*, 40(3), 540–561. <https://doi.org/10.1093/applin/amx007>
- Mauranen, A. (2020). Good texts in non-standard English: ELF and academic writing. In *ELF Research Methods and Approaches to Data and Analyses* (pp. 57-80). Routledge.
- Mauranen, A. (2012). *Exploring ELF: Academic English shaped by non-native speakers*. Cambridge University Press.
- McEnery, T., & Hardie, A. (2012). *Corpus Linguistics*. Cambridge University Press. <https://doi.org/10.1017/cbo9780511981395>
- Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? *Research in Nursing & Health*, 29(5), 489–497. <https://doi.org/10.1002/nur.20147>
- Pradita, I., & Puspitasari, D. A. (2025). The PRAYA Corpus: A Specialized Collection of Indonesian-Authoried ELT Research Articles. In *International conference on WorldS4* (pp. 243-251). Springer, Cham.

- Strauss, P., & Manalo, E. (2023). "Most good papers are published in English" Japanese academics' perspectives on the benefits and drawbacks of writing and publishing in English. *Journal of English for Research Publication Purposes*, 4(1), 5-28.
- Suhandoko, Ningrum, D. R., & Mustofa, A. (2025). A corpus-based study of authorial presence in the academic writing of Indonesian higher education EFL learners. *Studies in English Language and Education*, 12(3), 1246-1263.
- Taguchi, N. (2009). Corpus-informed assessment of comprehension of conversational implicatures in L2 English. *TESOL quarterly*, 43(4), 738-749.
- Triyoko, H., Wijana, I. D. P., & Baryadi, I. P. (2021). Hedges and boosters in Indonesian scientific articles. *Register Journal*, 14(1), 65-82.
- Wray, A. (2013). Formulaic language. *Language teaching*, 46(3), 316-334.
- Xi, X. (2017). What does corpus linguistics have to offer to language assessment? *Language Testing*, 34(4), 565-577.
- Yilmaz, S., & Römer, U. (2020). A corpus-based exploration of constructions in written academic English as a lingua franca. *Advances in Corpus-based Research on Academic Writing: Effects of discipline, register, and writer expertise*, 95, 59.
- Zhao, N. (2023). A corpus-based comparison study of first- person pronoun we in English-language abstracts. *Journal of English for Academic Purposes*, 63, 101244.