

# Formality in the Academic Writing of Thai EFL English-Major Students

**PONG-AMPAI KONGCHAROEN**

*Faculty of Humanities, Kasetsart University, Thailand*

**JIRAPORN DHANARATTIGANNON**

*Faculty of Humanities, Kasetsart University, Thailand*

**INTIRA BUMRUNGSALEE\***

*Faculty of Humanities, Kasetsart University, Thailand*

**Corresponding author email: [intira.b@ku.th](mailto:intira.b@ku.th)**

Article information	Abstract
<b>Article history:</b> Received: 19 Jun 2024 Accepted: 5 Aug 2024 Available online: 18 Apr 2025	<i>In recent years, there has been a growing trend of using informal styles in academic writing, including research articles. To examine the degree of formality in students' writing, this corpus-based study aimed to analyze the formal linguistic features in the academic writing assignments of English-major students at a Thai university. The learner corpus consisted of 552 assignments, totaling 190,506 words, and was organized into five different writing patterns. TagAnt was used to identify the part of speech for each word, while the Google Colab program was utilized for frequency counting. To assess the level of formality of the corpus, the F-score method proposed by Heylighen and Dewaele (1999) was applied. The results revealed that nouns were the most frequently used formal linguistic feature in student's essays. Despite this, the formality score of the learner corpus ranged from 51 to 53 across five years, indicating a relatively low level of formality. This score suggests that the students' writing was only slightly more formal, signaling the need for further development in grammatical complexity to achieve higher formality in their work. The results highlight the ongoing need for EFL teachers to instruct learners on formal linguistic features in academic writing.</i>
<b>Keywords:</b> Formality Corpus-based study F-score EFL students Academic writing	

## INTRODUCTION

Formality is an important stylistic characteristic of academic writing. As Chang and Swales (1999), and Swales and Feak (2012) highlighted, academic writing is expected to be in a formal style. When teaching academic writing to second language learners, formality seems to be emphasized and L2 learners are instructed to avoid informal features such as first-person pronouns, pronominal anaphoric references, and sentence-initial conjunctions or conjunctive adverbs. Nonetheless, language gradually changes, including writing styles. Many studies have investigated language features expressing formality and informality, including Baleghizadeh and Asadi (2022), Chang and Swales (1999), Eder et al. (2023), Heylighen and Dewaele (1999), Hyland and Jiang (2017), and Tocalo et al. (2022). Other studies have sought to explore changes

in formal styles of academic writing, finding that there has been a shift from formal to informal writing styles (Chang & Swales, 1999; Hyland & Jiang, 2017; Swales & Feak, 2012). The corpora used in these studies were mostly research articles written by native and non-native English speakers. There are a few studies conducted on informal and formal styles of general writing, such as emails, text messages, and student essays. Among them, Hyland and Jiang (2017) found a small increase in the use of informal features in research articles. Hyland and Jiang (2017) discussed this change in that “[t]hey also change for reasons which more directly relate to the rhetorical purposes of the genre and gradual adjustments to norms of interpersonal persuasion rather than efforts to weaken existing structures in favor of more ‘friendly’, ‘relaxed’ or conversational practices” (p. 49). Hyland and Jiang (2017) also raised the issue of L2 learners when they write English essays, since gradual rhetorical shift from a formal to an informal style, it is difficult for students and novice writers, particularly L2 students who are informed to write academic essays in a formal style. This is evident in Chang and Swales’ (1999) with the result that the majority of the L2 graduates in their study felt uneasy when informality was allowed because they followed the rules of formal academic English writing and combining formal and informal language in their academic writing effectively can be difficult and confusing for them.

In Thailand, formal academic writing has been taught for many years. Although the shift from a formal to an informal style has increased, students are still familiar with the formal style of writing and are instructed to use an appropriate academic writing style (Prapobratanakul, 2024) rather than informal features in their academic writing because it should be precise, concise, and unambiguous. Heylighen and Dewaele (1999) argued that formality can assist in avoiding ambiguity in an expression. Since few studies have been conducted on formality and informality among tertiary L2 students, especially in the Thai context, it is worth exploring how Thai students use formality features in their academic writing. Moreover, most previous studies focused on research articles, which inherently use formal writing. The present study instead investigates students’ written essay assignments in an academic writing course to identify how students, after being instructed on academic writing and formal writing styles, use formal linguistic features to avoid informal features when writing an academic essay at university. Although this writing course focuses on process-based writing which seeks to encourage students to create their voice (identity) in their pieces of writing, it is interesting to explore the extent to which the students use formal language features in their writing assignments.

## **Research questions**

This study addressed the following research question:

1. How did EFL English-major students use formality in terms of linguistic features in their writing assignments?
2. Are there any factors affecting students’ use of formal language features in their essay writing?

## LITERATURE REVIEW

### Formality

Liardét et al. (2019) state that English for Academic Purposes (EAP) students tend to be concerned about formality when doing academic writing tasks. This is also true for EFL learners when they write academic writing. They appear to employ the written mode using a greater degree of formality such as avoiding spoken language and informal words. Liardét et al. (2019) also observed that when defining formality, researchers often compare formal expressions in terms of contrasting characteristics of informal styles like personal, subjective reference, and spoken language.

Other scholars such as Howard (2014) and Hyland and Jiang (2017) noticed that the construct of formality in academic writing tends to be taught by focusing on the subject and audience, with few studies focusing on linguistic features.

Some academic writing textbooks often list linguistic features that learners should avoid in order to write with a formal tone and style. These linguistic features include using clear, concise, discipline-specific vocabulary while avoiding strongly emotive language, slang, colloquial expressions, and phrasal verbs. Additionally, academic writing discourages the use of personal pronouns such as 'you' and 'we', as well as abbreviations and contractions.

Morley-Warner (2010 as cited in Liardét et al., 2019) suggested that formality in writing can be achieved through the combination of several linguistic features. These features include structuring sentences so that the initial focus is on the issue rather than the writer and choosing pronouns that position the reader at a distance. Precision in verb choice, occasional use of passive construction, and the prevalence of abstract nouns contribute to a formal tone. Moreover, technical terms should be used where relevant, and nominalization can be employed when the text emphasizes ideas or concepts rather than actions. Finally, adopting a tentative tone can further enhance the formality of academic writing.

In their study on *Defining formality: Adapting to the abstract demands of academic discourse*, Liardét et al. (2019) concluded "in genres such as those examined in this study, formal expression tends to focus on concepts and how they relate (i.e., through definition, comparison and evaluation) rather than foregrounding human entities who impose their beliefs or perspectives toward these concepts (i.e., through personal statements, author-attributed evaluations and modulation)" (p. 157).

A few empirical studies have explored formality in learners' academic writing in different dimensions of variation, particularly on context-independent and precision, namely, the analysis factors used for measuring is "a measure of the 'complexity' of sentence structures, independent of their degree of formality" (Heylighen & Dewaele, 1999, p. 22).

## Formality score

The formula to calculate formality score or F-score in languages was first introduced by Heylighen and Dewaele (1999). Formality can be broken down into two main categories: surface formality and deep formality. Surface formality pays attention to form itself, while deep formality focuses on the unequivocal understanding of the form used. In other words, deep formality refers to the clarity of expression achieved by minimizing ambiguity through explicit and precise contextual descriptions. In the present paper, aligned with Heylighen and Dewaele (1999), deep formality is used because deep formality can be measured by forms of the writing style along with the writer's original intention or purpose that cannot reflect in the surface formality.

A formal writing style is characterized by detachment, accuracy, rigidity, and weightiness, while an informal style tends to be more flexible, direct, implicit, and engaging, albeit less informative. To empirically measure formality, the F-score is suggested, which is based on the distribution of word classes or word forms in a given corpus. Nouns, adjectives, articles, and prepositions are more prevalent in formal writing, whereas informal writing tends to feature a higher frequency of pronouns, adverbs, verbs, and interjections (Heylighen & Dewaele, 1999).

From F-score proposal, the noun modification proposed by Biber (1988 as cited in Jitpraneechai, 2019) can be an indicator of formality in a written text. Nouns, adjectives, articles, and prepositions are the structures of noun phrases. Biber (1988 as cited in Jitpraneechai, 2019) proposed that in formal language, phrases are used more frequently than clauses. The more frequent the phrases are, the more formal the language is. Subsequently, the noun phrases indicate the formal writing style.

The main criticism of F-score measure is that their measure focuses only on the frequencies of word categories and ignores other stylistic variations such as situation and interlocutors. However, Heylighen and Dewaele (1999) argue that their measure "seems to do its job, unambiguously distinguishing types of language which we would intuitively and theoretically expect to differ in formality. The advantage of such a coarse-grained approach is that it facilitates the collection and processing of data for different samples or styles" (p. 24). Therefore, they propose that F-score is a reliable measure of formality.

Heylighen and Dewaele (1999) proposed the following formula for formality (F-score):

$$F = (\text{noun frequency} + \text{adjective frequency} + \text{preposition frequency} + \text{article frequency} - \text{pronoun frequency} - \text{verb frequency} - \text{adverb frequency} - \text{interjection frequency} + 100)/2$$

From the F-score formula, nouns, adjectives, and prepositions are classified as positive function words because they are context-independent. In contrast, pronouns, adverbs, and interjections fall under the negative function, as they are context-dependent. The formula produces percentages, which indicate how frequently words from each category appear in relation to the total word count. The formality score (F-score) ranges from 0 to 100%, with a higher score

representing more formal language. In Dutch scientific texts, the F-score averaged around 66, while in newspapers, it was slightly higher at 68. Magazines scored 58 and novels had a lower F-score of 52. In spoken genres, the F-score of academic speech was 44, while non-academic speech scored 40. Additionally, similar F-scores were found in Italian dictionaries studied by Bortolini et al. (1971 as cited in Heylighen & Dewaele, 1999), and Juilland and Traversa (1973 as cited in Heylighen & Dewaele, 1999), confirming the consistency of these results across different languages. The data from them also suggests that written discourse tends to be more formal than spoken discourse though no precise cut-off point can definitively indicate formality. Therefore, the F-score is reported as a percentage where a higher percentage reflects greater formality (Heylighen & Dewaele, 1999).

Many studies investigating formality in communication often focus on the rise of informality, as it has been increasingly observed in academic texts. For example, Hyland and Jiang (2017) noted a shift from formal to informal styles in research reports and other academic texts. Some studies compare the use of formality and informality in both formal and informal texts to understand this shift. In this study, we examine the formal linguistic features in students' written assignments for their first academic writing course. These students had just been introduced to formal academic writing. The F-score is particularly suitable for this analysis because it offers an objective, mathematical approach that has been used in many studies on formality and informality. This section reviews existing research that employed the F-score for data analysis, focusing on its use in assessing formality in written texts.

## **Related studies**

### **Formality in EFL learners academic writing: A cross-cultural perspective**

Akhtar and Riaz (2019) explored academic writing among undergraduate EFL learners by analyzing essays from students at the Government Postgraduate College for Women in Pakistan. The researchers found that students predominantly used formal language but exhibited numerous linguistic errors. Despite the high formality level, these errors reflect limitations in language proficiency. The academic environment and limited exposure to informal English were seen as major contributors. This study highlights the importance of interactive teaching methods aimed at improving rhetorical and stylistic competence, calling for focused interventions in EFL teaching. It also encourages further research into formality levels across different academic texts, suggesting a need for targeted interventions to improve stylistic competence among undergraduate EFL learners.

Praminatih et al. (2018) examined the decline of informal language features in Indonesian undergraduate thesis abstracts. The researchers found a shift toward more formal writing, with students increasingly avoiding informal elements like sentence fragments and direct questions. This shift is possibly due to improvements in teaching methods or evolving academic norms. The study demonstrates that exposure to formal writing standards can lead to significant improvements in student writing, a key point for EFL instructors.

Baleghizadeh and Asadi (2022) compared formality in research articles written by Iranian non-native researchers and native English speakers. The findings indicated a higher level of formality in the writings of Iranian researchers, likely influenced by strict academic traditions in Iranian education. The heightened formality was linked to challenges in global academic publishing, as excessive formality might obscure the author's stance. The study underscores the need for balanced instruction in formality to help students navigate both local and international academic standards. Additionally, this study highlights potential consequences of this heightened formality, including possible rejections by international journal editors due to a lack of authorial stance on the results. An implication for academic writing and international scientific communication is thus required.

### **Formality across genres**

Eriksson (2017) explored formality levels in sports journalism, revealing that the type of sports, for example, horse polo versus soccer, influenced the degree of formality in writing. The study focused on two distinct types of newspapers: a broadsheet (The Daily Telegraph) and a tabloid (The Daily Express). Horse polo articles from The Daily Telegraph had a formality score of 66.07, while those from The Daily Express scored 65.85, while soccer articles scored lower, with The Daily Telegraph at 56.85 and The Daily Express at 56.32. This indicated that articles on horse polo were more formal and scored higher on the F-score measure than those on soccer, even when published in both broadsheet and tabloid newspapers. Although horse polo articles typically exhibited features associated with informational production, characterized by a higher ratio of nouns, long words, and adjectives, this style is often found in more formal writing, such as academic papers and legal documents. Soccer articles, in contrast, followed a more involved impersonal production style, characterized by a higher ratio of verbs, adverbs, pronouns, and WH-questions. This style is more typical of spoken interaction and less formal writing. This study suggests that while genre plays a role in determining formality, audience expectations also shape the language used, a consideration that EFL learners should be mindful of in academic writing.

Karlsson (2008) examined the formality levels of websites in the U.S. and U.K. The study found that the key linguistic features that define formality in websites included lexical formality, syntactic complexity, and register. While American websites were generally more formal than their British counterparts, the level of formality varied by the type of content, for example, "About Us" pages versus "Jobs/Careers" pages. The American sites scored higher on the F-score measure with the score of 70.2 in the "About Us" web pages. On the contrary, the formality score of the British "About Us" is 64.4. In addition, the difference in formality is less pronounced in the "Jobs/Careers" web pages, with the formality score of 61.74 in the American "Jobs/Careers" webpages and 59.5 in British "Jobs/Careers" webpages, indicating that the context of the content may influence the level of formality. These findings suggest that even in digital contexts, formality remains an important aspect of written communication, and EFL students should learn to adapt their writing to fit different genres and platforms. However, the findings indicate that while there is a connection between the level of formality and the country of origin, this connection is somewhat vague and not absolute. Regarding market sector, there was no clear connection between the level of formality in websites and the

market sector. This suggests that factors other than market positioning may play a more significant role in determining language formality.

Keppens (2015) highlighted the varying degrees of formality in different written genres, from academic papers, conference papers, TED talks and popular science articles. Academic papers displayed the highest formality (score of 69.95), followed by conference papers of 69.29, while popular science texts were ranked the lowest formality score. This variation suggests that students need to be aware of the specific formality requirements of each genre, reinforcing the importance of instruction in genre-based formality.

### **Formality and individual differences**

Nowson et al. (2005) explored the impact of gender and personality on formality in weblogs. The researchers asked bloggers to complete a sociobiographic questionnaire and an online implementation of an IPIP Five Factor Personality Inventory. Each blogger was also requested to submit text they had previously written over the month preceding the questionnaire to reduce the effects of subjects choosing their best or favorite month. The study found variability in the formality of weblogs due to individual differences among writers. Males generally preferred a more formal writing style, while females used a more contextual style, with females scoring lower on the F-score measure, indicating a preference for more contextual language compared to males. Personality traits like Agreeableness and Openness also played a significant role in determining formality levels than Extraversion and Neuroticism. While blogs are less formal than academic writing, blogs are generally more formal than emails but less formal than biographies with the formality score of 55.2, 53.1, and 56.3 respectively. This positions blogs as a unique genre that is relatively informal compared to traditional written forms like academic essays. This research illustrates how individual differences can influence writing style, a factor that EFL teachers should consider when guiding students toward achieving the appropriate level of formality in their academic work.

### **The need for formality instruction in EFL writing**

Despite the evolving trends in academic writing, where informality may be more acceptable in certain contexts, formal writing remains a critical skill for EFL students, particularly in academic and research settings. For Thai EFL students, understanding and practicing formal language features is vital for academic success. Therefore, this research underscores the necessity for EFL teachers to continue emphasizing formal writing conventions. While students should be encouraged to explore formal language features and adapt those to different genres, instruction in formal linguistic features should remain central to EFL curricula in order to ensure that students are prepared to meet the challenges of both local and international academic standards.



## METHODOLOGY

### Data collection

The learner corpus utilized in this research comprised five types of writing tasks from an English writing course over five academic years between 2017 and 2021. A breakdown of the data for each academic year is presented below.

- In 2017, 38 students completed four assignments (Listing, Sequence, Comparison and Contrast, Cause-effect), totaling 152 assignments and 50,496 words.
- In 2018, 18 students completed the same four types of assignments, totaling 72 assignments and 21,470 words.
- In 2019, 25 students completed the same four types of assignments, totaling 100 assignments and 30,827 words.
- In 2020, 22 students completed the same four types of assignments, totaling 88 assignments and 27,938 words.
- In 2021, 35 students completed four types of assignments (Listing, Comparison and Contrast, Cause-effect, and Problem-solution), totaling 140 assignments and 59,760 words.

In total, the learner corpus included 552 assignments comprising approximately 190,506 words. In 2021, the new Revised Curriculum 2020 was used which revised the writing course in terms of the organizational patterns taught. This change involved replacing the Sequence assignment with the Problem-solution assignment. Although the total token of Problem-Solution assignment was less than that of other paragraph types, the number of each grammatical feature had been normalized before entering the formula.

The data in this research was approved by the Institutional Review Board (IRB), Kasetsart University Research Ethics Committee, on November 21, 2022 (COE No. COE65/147).

### Data analysis

First, TagAnt was used to identify the part of speech of each word. TagAnt is a free program that can generate the part of speech taggers in the corpus. TagAnt is a straightforward tagging tool which utilizes the TreeTagger engine, created by Schmid (1994). It functions by processing either a single input text or a batch of text files, and then applies Parts-Of-Speech (POS) tags to them. TagAnt was developed using Python and Qt, with the PyInstaller compiler used to create executables for various operating systems, and is available at <https://www.laurence-anthony.net/software/tagant/> (Anthony, 2024). The tagger from TagAnt generated the following parts of speech.



**Table 1**  
**Parts of speech used in tagger from TagAnt**

POS Tag	Description	Example	POS Tag	Description	Example
CC	coordinating conjunction	<i>and, but, or, &amp;</i>	VB	verb <i>be</i> , base form	<i>be</i>
CD	cardinal number	<i>1, three</i>	VBD	verb <i>be</i> , past	<i>was / were</i>
DT	determiner	<i>the</i>	VBG	verb <i>be</i> , gerund/participle	<i>being</i>
EX	existential there	<i>there is</i>	VBN	verb <i>be</i> , past participle	<i>been</i>
FW	foreign word	<i>d'œuvre</i>	VBZ	verb <i>be</i> , pres, 3rd p. sing	<i>is</i>
IN	preposition/subord. conj.	<i>in, of, like, after, whether</i>	VBP	verb <i>be</i> , pres non-3rd p.	<i>am / are</i>
IN/that	complementizer	<i>that</i>	VD	verb <i>do</i> , base form	<i>do</i>
JJ	adjective	<i>green</i>	VDD	verb <i>do</i> , past	<i>did</i>
JJR	adjective, comparative	<i>greener</i>	VDG	verb <i>do</i> gerund/participle	<i>doing</i>
JJS	adjective, superlative	<i>greenest</i>	VDN	verb <i>do</i> , past participle	<i>done</i>
LS	list marker	<i>(1),</i>	VDZ	verb <i>do</i> , pres, 3rd per.sing	<i>does</i>
MD	modal	<i>could, will</i>	VDP	verb <i>do</i> , pres, non-3rd per.	<i>do</i>
NN	noun, singular or mass	<i>table</i>	VH	verb <i>have</i> , base form	<i>have</i>
NNS	noun plural	<i>tables</i>	VHD	verb <i>have</i> , past	<i>had</i>
NP	proper noun, singular	<i>John</i>	VHG	verb <i>have</i> , gerund/participle	<i>having</i>
NPS	proper noun, plural	<i>Vikings</i>	VHN	verb <i>have</i> , past participle	<i>had</i>
PDT	predeterminer	<i>both the boys</i>	VHZ	verb <i>have</i> , pres 3rd per.sing	<i>has</i>
POS	possessive ending	<i>friend's</i>	VHP	verb <i>have</i> , pres non-3rd per.	<i>have</i>
PP	personal pronoun	<i>I, he, it</i>	VV	verb, base form	<i>take</i>
PP\$	possessive pronoun	<i>my, his</i>	VVD	verb, past tense	<i>took</i>
RB	adverb	<i>however, usually, here, not</i>	VVG	verb, gerund/participle	<i>taking</i>
RBR	adverb, comparative	<i>better</i>	VVN	verb, past participle	<i>taken</i>
RBS	adverb, superlative	<i>best</i>	VVP	verb, present, non-3rd p.	<i>take</i>
RP	particle	<i>give up</i>	VVZ	verb, present 3d p. sing.	<i>takes</i>
SENT	end punctuation	<i>?, !, .</i>	WDT	wh-determiner	<i>which</i>
SYM	symbol	<i>@, +, *, ^, /, =</i>	WP	wh-pronoun	<i>who, what</i>
TO	to	<i>to go, to him</i>	WP\$	possessive wh-pronoun	<i>whose</i>
UH	interjection	<i>uhhuhhuhh</i>	WRB	wh-abverb	<i>where, when</i>
			:	general joiner	<i>;, ~, --</i>
			\$	currency symbol	<i>\$, £</i>

Each part of speech was counted with the assistance of Google Colab, a program that can be activated by simple code writing to command the program to categorize or count the particular words or word forms. As stated on the website itself, Colab is a Jupyter Notebook service that is hosted, eliminating the need for setup, and offers complimentary access to computational assets like GPUs and TPUs. It is particularly advantageous for tasks related to machine learning, data science, and educational purposes (Google Colaboratory, n.d.)

To ensure accuracy of the formula, normalization is required. All frequencies are put into per thousand words. For example, if the counted frequency is 238, the normalized score is 0.238. After that, the F-score formula was used to calculate the formality of the corpus.

## RESULTS

This section provides a comprehensive breakdown of the frequency of grammatical categories used by EFL English-major students in their writing assignments. The results from TagAnt and Google Colab reveal how Thai EFL English-Major students who were the participants in this study used formal linguistic features in their writing assignments. Table 2 presents frequency of grammatical categories students used.

**Table 2**  
**Frequency of grammatical categories students used**

Year	Organizational patterns	Total frequency	Articles	Prepositions	Adjectives	Nouns	Adverbs	Interjection	Verbs	Pronouns
2017	Listing	105078	592	1167	871	2367	922	5	2613	60
2017	Sequence	106518	753	1280	712	2297	991	6	2719	45
2017	Cause&Effect	131086	734	1604	1174	3102	1211	6	3061	102
2017	Compare&Contrast	151758	1064	2064	1223	3598	1326	3	3649	70
2018	Listing	33674	174	362	293	755	288	3	914	15
2018	Sequence	33748	213	392	221	704	358	5	920	16
2018	Cause&Effect	58698	358	755	476	1522	596	4	1307	24
2018	Compare&Contrast	67898	697	857	584	2063	457	2	1295	38
2019	Listing	55219	298	637	451	1367	472	0	1374	28
2019	Sequence	56030	364	674	434	1179	533	5	1480	25
2019	Cause&Effect	91108	635	1232	850	2477	735	0	1985	37
2019	Compare&Contrast	88152	592	1192	789	2398	728	5	1968	36
2020	Listing	45319	258	491	372	990	436	4	1227	19
2020	Sequence	52185	397	601	411	1102	539	9	1374	26
2020	Cause&Effect	95956	689	1165	923	2570	770	5	2134	47
2020	Compare&Contrast	113075	914	1451	966	3128	867	2	2331	35
2021	Listing	138687	878	1602	1230	3310	183	11	3455	97
2021	Problem&Solution	143588	910	1648	1311	3732	1132	5	3512	48
2021	Cause&Effect	142332	868	1822	1301	3950	1092	6	3333	82
2021	Compare&Contrast	144525	1139	1847	1320	3868	2019	1	3184	55
Total			12527	22843	15912	46479	16655	87	43835	905

Table 2 shows the raw frequency counts of various grammatical categories used by the participants and breakdown of articles, prepositions, adjectives, nouns, adverbs, interjections, verbs, and pronouns across different organizational patterns and years (2017-2021). Nouns and verbs appear to be the most frequently used grammatical category, with the frequency of 46479 and 43835, respectively.

Prepositions, which also play a significant role in students' writing, suggest that students use a variety of complex structures. They are slightly more common with the frequency of 22843. Adjectives and adverbs are used at similar rates, with frequencies of 15912 and 16655, respectively. In contrast, interjections, which signify informality, are used minimally, with the frequency of 87.

**Table 3**  
**Frequency of grammatical categories that show formality**

Year	Organizational patterns	Articles	Prepositions	Adjectives	Nouns
2017	Listing	592	1167	871	2367
2017	Sequence	753	1280	712	2297
2017	Cause&Effect	734	1604	1174	3102
2017	Compare&Contrast	1064	2064	1223	3598
2018	Listing	174	362	293	755
2018	Sequence	213	392	221	704
2018	Cause&Effect	358	755	476	1522
2018	Compare&Contrast	697	857	584	2063

Year	Organizational patterns	Articles	Prepositions	Adjectives	Nouns
2019	Listing	298	637	451	1367
2019	Sequence	364	674	434	1179
2019	Cause&Effect	635	1232	850	2477
2019	Compare&Contrast	592	1192	789	2398
2020	Listing	258	491	372	990
2020	Sequence	397	601	411	1102
2020	Cause&Effect	689	1165	923	2570
2020	Compare&Contrast	914	1451	966	3128
2021	Listing	878	1602	1230	3310
2021	Problem&Solution	910	1648	1311	3732
2021	Cause&Effect	868	1822	1301	3950
2021	Compare&Contrast	1139	1847	1320	3868
	Total	12527	22843	15912	46479

Table 3 illustrates the grammatical categories that indicate formality according to Heylighen and Dewaele's (1999) F-score formula. The table narrows down the grammatical features contributing to formal writing, primarily focusing on nouns, adjectives, articles, and prepositions. When examining the positive or formal linguistic features outlined by Heylighen and Dewaele's formula, the total frequency of all grammatical categories is 97761. In other words, these formal linguistic features are commonly found in the students' writing.

Nouns remain the most prominent feature across all years and organizational patterns, contributing greatly to formality. This is consistent with Heylighen and Dewaele's (1999) formality model, where nouns represent context-independent words crucial for formal writing. This tendency to use noun phrases also highlights the formality of their writing style. This observation aligns with Biber's (1988 as cited in Jitpraneechai, 2019) affirmation that phrases are more commonly used than clauses in formal language. Moreover, adjectives are also frequently used, contributing to descriptive precision in academic writing. Prepositions similarly play a critical role in maintaining the formal structure, as they help articulate relationships between different entities in the text.

**Table 4**  
**Frequency of grammatical categories that show informality**

Year	Organizational patterns	Pronouns	Adverbs	Verbs	Interjections
2017	Listing	60	922	2613	5
2017	Sequence	45	991	2719	6
2017	Cause&Effect	102	1211	3061	6
2017	Compare&Contrast	70	1326	3649	3
2018	Listing	15	288	914	3
2018	Sequence	16	358	920	5
2018	Cause&Effect	24	596	1307	4
2018	Compare&Contrast	38	457	1295	2
2019	Listing	28	472	1374	0
2019	Sequence	25	533	1480	5
2019	Cause&Effect	37	735	1985	0
2019	Compare&Contrast	36	728	1968	5

Year	Organizational patterns	Pronouns	Adverbs	Verbs	Interjections
2020	Listing	19	436	1227	4
2020	Sequence	26	539	1374	9
2020	Cause&Effect	47	770	2134	5
2020	Compare&Contrast	35	867	2331	2
2021	Listing	97	183	3455	11
2021	Problem&Solution	48	1132	3512	5
2021	Cause&Effect	82	1092	3333	6
2021	Compare&Contrast	55	2019	3184	1
	Total	905	16655	43835	87

Table 4 presents the grammatical categories associated with informality according to Heylighen and Dewaele's (1999) F-score formula. Adverbs, pronouns, and verbs appear more frequently in informal writing, as reflected in Table 4. Pronouns are particularly low across the years, which is expected in academic writing, where formal nouns and structures are preferred over personal references. The use of verbs is still high, but as verbs are probably necessary for action-oriented academic writing, it reflects students' ability to express ideas dynamically, even in formal assignments. Nevertheless, verbs do lean more towards informality. In addition, adverbs, often used to modify verbs, appear less frequently than nouns and adjectives but are still present in varying degrees across years and patterns. When considering the negative or informal linguistic features from Heylighen and Dewaele's formula, the total frequency of all grammatical categories is 61842. This number is lower than the total frequency found in Table 2. This suggests that the students in this study use informal linguistic features less frequently than formal ones.

### Formality in L2 Thai university students' essay writing based on organizational patterns from 2017 to 2021

Table 5  
F-score results

Year/Organizational patterns	2017	2018	2019	2020	2021
Listing	51.397	50.364	50.879	50.424	52.274
Sequence	51.281	50.231	50.608	50.563	N/A
Compare & Contrast	52.9	52.409	52.233	53.224	52.915
Cause & Effect	52.234	51.18	52.436	52.39	53.427
Problem-Solution	N/A	N/A	N/A	N/A	52.904

Table 5 presents the formality score of the learner corpus categorized by the organizational patterns. The table presents F-scores (formality scores) for different organizational patterns in EFL Thai university students' essay writing from 2017 to 2021. The F-score, based on Heylighen and Dewaele's (1999) model, quantifies the level of formality in writing by calculating the frequency of context-independent words (like nouns, prepositions) compared to context-dependent ones (like pronouns, adverbs). An F-score closer to 100 indicates a higher level of formality. The Sequence shows the least formal F-score was 50.231 in 2018, while the highest formality falls into Cause and Effect in 2021 with an F-score of 53.427. When looking at the

whole organizational patterns, Listing and Sequence appear less formal in writing styles than the other three organizational patterns. Across the years, Compare and Contrast and Cause and Effect seem to have equivalent F-scores at about 52 to 53. Despite the decreased F-score in Cause and Effect in 2018, the rest of the formality seems to be high compared with Listing and Sequence. There is slight fluctuation across the organizational patterns and years, but generally, the formality level stays consistent, with the most common organizational patterns, for example, Cause and Effect, Compare and Contrast, demonstrating higher formality levels. This may be because the language structures of Compare and Contrast and Cause and Effect tend to use many noun phrases which can be a formal indicator.

The pattern of formality in L2 Thai university students' essay writing from 2017 to 2021 can be described as follows.

In 2017, the Compare and Contrast organizational pattern had the highest formality score, indicating that essays using this pattern were the most formal. Conversely, Sequence had the lowest formality score, indicating less formality. In 2018, Compare and Contrast again had the highest formality score, while Sequence had the lowest score, indicating a consistent trend. In 2019, the Cause and Effect pattern had the highest formality score, closely followed by Compare and Contrast. Sequence had the lowest score, indicating a decrease in formality compared to the previous years.

In 2020, Compare and Contrast had the highest formality score, showing a significant increase in formality compared to previous years. Listing had the lowest score, indicating a decrease in formality compared to Cause and Effect and Compare and Contrast. In 2021, the Cause and Effect pattern had the highest formality score, and a new pattern, Problem and Solution, was introduced with a high formality score. Listing had the lowest score, but it was still higher than previous years, indicating an overall increase in formality.

The data indicates that the pattern of formality in L2 university students' essay writing from 2017 to 2021 in this study changed over time. The data reveals that Compare and Contrast consistently ranked among the highest in formality, with a peak in 2020. In contrast, Sequence had the lowest scores in 2017 and 2018, indicating less formality. Interestingly, Listing had the lowest scores in 2019 and 2020, while Cause and Effect achieved the highest score in 2021.

As Table 3 portrays, the linguistic feature of formality that the students used the most was nouns, followed by adjectives and prepositions. It is also noted that the students used informal linguistic features in a slightly smaller amount than the formal linguistic features and the most common informal feature used was verbs followed by adverbs (Table 4). This is not surprising because these main features of both formality and informality are the main parts of a sentence structure.

Overall, the results show that EFL English-major students in this study largely align with formal linguistic expectations in their writing assignments. The high frequency of nouns, prepositions, and adjectives reflects a focus on objective, descriptive, and informative writing, consistent with academic standards. These students exhibit a good understanding of formal grammatical

features, with minimal use of pronouns, interjections, and adverbs, which are typically associated with informality. However, despite this focus on formality, the results still show room for improvement. The formality scores in the data indicate that while students are incorporating formal features, the presence of verbs and adverbs suggests a level of informality still present in their writing. This is reflective of the ongoing trend noted in the literature, such as Hyland and Jiang's (2017) work, where there is an increasing shift towards a blend of formal and informal styles in academic texts. In addition, there was a noticeable increase in the formality of essays across various organizational patterns over the years, with 2021 showing the highest formality scores. The data indicates that the formality of L2 Thai university students' essay writing increased over the five years, with certain organizational patterns such as Compare and Contrast and Cause and Effect consistently leading to more formal writing.

## DISCUSSION

### **How did EFL English-major students use formality in terms of linguistic features in their writing assignments?**

In response to the research question, Thai EFL English-major students in this study used formal linguistic features slightly higher than informal ones. As suggested by Heylighen and Dewaele (1999), formal writing tends to feature context-independent words such as nouns and adjectives, while informal writing favors pronouns and adverbs. The data from this study support this, as students predominantly use nouns and adjectives, indicating that they can use appropriate formal linguistic features in their academic writing. When comparing the F-scores with previous studies, their writing has formality in a low rate with an F-score of 51-53 while a more formal piece of writing in those studies reveals the F-score over 60. This can be attributed to the nature of their writing course. This course is the first academic writing class for second-year English-major students at the university. In school, Thai EFL students rarely engaged in academic writing, often focusing on narrative and descriptive styles, without much emphasis on formal writing. At the university level, however, this academic writing course emphasizes formal styles. The students were taught to use clear, concise, discipline-specific vocabulary, avoid strongly emotive language, and steer clear of slang and colloquial expressions. The course also followed the process-based approach, where students learned essential writing components such as developing a voice, writing with purpose and audience in mind, and creating authorship. Praminatih et al. (2018) discussed the importance of process-oriented instruction, which includes stages like drafting, revising, and receiving feedback. This approach helps students develop a more refined writing style. Praminatih et al. (2018) also noted that explicit teaching of formal writing conventions, such as avoiding emotive language and slang, can influence students to adhere to a more formal writing style. Therefore, the design of the university academic writing course likely contributes to the Thai students' preference for formal language, as reflected in the F-scores. This also indicates a significant opportunity to develop Thai EFL English major students' writing competence in terms of grammatical complexity, allowing them to achieve greater formality in their writing.



In terms of formal linguistic features used, the students demonstrated an emerging proficiency in formal writing through the use of nouns. However, they occasionally blended formal features with informal ones, particularly in the use of verbs. The possible explanation is that the sentences they wrote comprised a simple noun phrase and a verb, adhering to the obligatory subject-verb pattern in sentence structure, resulting in the almost equal number of nouns and verbs used in their writing. The students struggled to incorporate complex noun phrases in their writing assignments, primarily using noun phrases in subject and object positions. This limitation likely stemmed from their lack of exposure to complex noun phrase structures, which would be thoroughly addressed in the next two English Structure courses they were scheduled to enroll.

While the students show a promising use of formal linguistic features, writing instruction in EFL contexts should continue to emphasize the importance of minimizing informal language and encourage more consistent use of formal elements in academic assignments. This will better prepare students to meet the high standards of formality required in professional academic writing.

### **Are there any factors affecting students' use of formal language features in their essay writing?**

The results suggest that organizational patterns (listing, sequence, comparison and contrast, cause and effect, and problem-solution) appear to influence the use of formal linguistic features in the students' writing assignments. When examining each organizational pattern, students used more formal styles in Compare and Contrast, Cause and Effect, and Problem and Solution essays. This may be due to the chronological order of assignments. Listing and Sequence were assigned first, before the midterm examination. Since students initially had less experience with academic writing and were accustomed to an informal style from school, they struggled to transition to a formal style. However, with F-scores around 50-51, they demonstrated a strong start in using formal style at the beginning of the course.

In addition, the essay writing organizational patterns of the English-major students from 2017 to 2021 show a consistent use of formal organizational structures, such as Compare and Contrast and Cause and Effect, particularly in more recent years. The shift towards more formal structures also corresponds with broader academic writing research findings. The formality patterns in the Thai L2 university students' essay writing can be linked to previous studies in other EFL contexts such as Akhtar and Riaz (2019) and Baleghizadeh and Asadi (2022), in that non-native English speakers of L2 learners tend to write research articles in a more formal style than native English speakers.

Moreover, students' essays that employ complex organizational patterns like Compare and Contrast and Cause and Effect tend to have higher F-scores. That means students attempt to use more formal language when tackling complex organizational patterns. These are in line with Akhtar and Riaz (2019) in that stylistic competence and the choice of linguistic features in EFL learners' writing impact formality, and advanced linguistic structures, which require more formal language, indicate a higher level of academic writing proficiency.



## CONCLUSION AND IMPLICATIONS

### Educational implications

#### *Formality with engaging writing*

Given the patterns of formality found in the Thai L2 university students' essay writing and the broader discussions within EFL research, it is important to consider whether teaching formality in Thai tertiary writing courses remains necessary.

According to Hyland and Jiang (2017), academic writing is becoming less formal due to the need for accessibility and engagement with a wider audience. The researchers noted an increase in informal features like contractions, first-person pronouns, and direct addresses to the reader in academic texts. Similarly, Tocalo et al. (2022) studied ESL academic writing in the Philippines, finding that while there has been an increase in informal features, the extent and type of informality vary significantly. For ESL writers in the Philippines, the increase in informality was primarily observed in the use of unattended anaphoric pronouns and sentence-initial conjunctions, while other features such as first-person pronouns and sentence-final prepositions were less common. This supports the idea that informality in academic writing is increasing. However, this does not affect all aspects of writing. To illustrate, although the informal style is gradually accepted in academic writing, the scholarly audience maintains some features and elements of formality when writing academic essays like research articles to be published in both national and international journals. This leads to a concern for teachers of English writing, particularly in ESL and EFL contexts, in regards to how to help their students balance formality and informality in their academic writing and to reach an acceptable standard of academic writing at higher levels, such as in research articles. One possible solution for this concern is that the teachers should study the formal and informal features commonly found in various academic texts and introduce their students those features and how they are used. Also, allowing students to examine the use of formal and informal features in authentic academic texts will help students be familiar with those features and understand how they are used so that they can apply them to their own academic writing.

#### Further suggestions

1. The results of this study correspond to the previous studies in that EFL learners tend to write their academic papers in formal and informal styles. However, formal language still has its place in academic writing (Chang & Swales, 1999; Eder et al., 2023; Heylighen & Dewaele, 1999; Tocalo et al., 2022). Studying formality in L2 learners will help us learn how the students write their academic essays, what features of formality and how they use them so that we can assist them to use appropriate language in academic writing based on their background while introducing them to informality. This suggests that the writing instructions and materials should change to meet this trend. In other words, for instructions, teachers should not focus solely on formal language, but they should allow an informal style in academic writing. The process-based approach used for the English writing course in this study can be a suitable approach to help students blend formality and informality without notice. The nature of this

approach invites both formal and informal styles to create a meaningful piece of writing with voice and authorship.

Moreover, teachers should teach students formal and informal linguistic features so they can use various linguistic features of both formal and informal styles when completing academic writing assignments. Teachers can include a corpus-based approach in their writing class to let the students engage in authentic language use and introduce them with specific text analysis software such as AntConc to help them learn about linguistic features of formality and informality to enable them to apply those to their writing. Additionally, authentic reading texts, particularly advanced ones like research articles, should be used so that students are aware of the conventions of English texts in different genres, and they can later employ those writing styles in their own academic writing in a more English-native speaker's.

2. Most studies on formality and informality in academic writing has been completed in the last decade, and most were conducted on the corpus of research articles rather than in other types of writing. More research is therefore necessary in this area to make a strong generalization about the trend of academic writing, changing from formal to informal in addition to the degree of change and how these changes are occurring. Moreover, previous studies focus only on the complexity of linguistic features. Other language features such as academic words should be studied to measure formality and informality.

3. Other factors are related to formality and informality in writing, for instance interlocutors, gender, and situations which should be included in future studies to understand how these factors affect the use of formality and informality in students' academic writing.

4. Studies on formality and informality should be conducted using L2 students' various academic writing assignments compared with periodical research articles to explore the features of formality and informality that are commonly used in those academic texts such as the use of attributive adjectives and nominalization for formality and the use of verbs, adverbs and the 1<sup>st</sup> pronouns for informality. The teachers can also introduce the common formal and informal features that are used in academic writing to L2 students in their academic writing class and help them use both formal and informal language appropriately.

## **Conclusion**

This study analyzed linguistic features of formality in the academic writing of English-major students. The results reveal that the students used more formal linguistic features than informal features. This suggests that since this is the first academic writing course offered to the English-major students, they tend to successfully combine formality and informality. They started using a formal style of writing more than an informal style. It also suggests that a process-based approach may be suitable as a fundamental approach for teaching EFL English-major students who have rarely experienced formal writing like academic writing. However, since the trend of academic writing is gradually changing and the informal style has become more accepted, writing instructors should be aware of this trend and improve their writing instruction by including informal styles of writing in their lessons as well.

## THE AUTHORS

**Pong-ampai Kongcharoen** is an assistant professor of English at the Department of Foreign Languages, Faculty of Humanities, Kasetsart University, Thailand. She also holds the position of EFL-IS (English as a Foreign Language Interest Section) Past Chair at TESOL International Association. Her research interests lie in corpus linguistics, semantics, discourse analysis, vocabulary learning and teaching, and second language acquisition.

[pongampai.k@ku.th](mailto:pongampai.k@ku.th)

**Jiraporn Dhanarattigannon**, PhD, is an assistant professor at the Department of Foreign Languages, Faculty of Humanities, Kasetsart University, Thailand. Her areas of interest include EFL writing (process-based approach), extensive reading, self-assessment, online language learning and teaching, and corpus-based research.

[jiraporn2000@yahoo.com](mailto:jiraporn2000@yahoo.com)

**Intira Bumrungsalee** is a lecturer at the Department of Foreign Languages, Faculty of Humanities, Kasetsart University, Thailand. She has a PhD in Translation Studies from Warwick University, UK and an MA in English from Chulalongkorn University, Thailand. Her academic interests include audiovisual translation, literary translation and cross-cultural communication.

[intira.b@ku.th](mailto:intira.b@ku.th)

## REFERENCES

- Akhtar, R., & Riaz, M. (2019). Formality in academic writing: Investigating stylistic competence of undergraduate EFL learners. *University of Wah Journal of Social Sciences*, 2(1), 1–20. <https://uwjss.org.pk/index.php/ojs3/article/view/59/34>
- Anthony, L. (2024). *TagAnt (Version 2.1.0)* [Computer Software]. Waseda University. <https://laurenceanthony.net/software.html>
- Baleghizadeh, S., & Asadi, F. (2022). A comparative study of text formality of Applied Linguistics articles written in English by Iranian and native speaking researchers. *Journal of Language Horizons*, 6(13), 7–25. <https://doi.org/10.22051/lghor.2021.30138.1472>
- Chang, Y., & Swales, J. (1999). Informal elements in English academic writing: Threats or opportunities for advanced non-native speakers? In C. Candlin & K. Hyland (Eds.), *Writing: Texts, processes and practices* (pp. 145–167). Longman.
- Eder, E., Krieg-Holz, U., & Wiegand, M. (2023) A question of style: A dataset for analyzing formality on different levels. *Findings of the Association for Computational Linguistics: EACL, 2023*, 580–593. <https://aclanthology.org/2023.findings-eacl.42.pdf>
- Eriksson, D. (2017). *Using the F-measure to test formality in sports reporting: A comparison of the language used in soccer and horse polo articles in two British newspapers* [Master's thesis, Karlstad University]. Karlstad University Library. <https://kau.diva-portal.org/smash/get/diva2:1223014/FULLTEXT01.pdf>
- Google Colaboratory. (n.d.). Google Colaboratory. <https://colab.google/>
- Heylighen, F., & Dewaele, J. (1999). *Formality of language: Definition, measurement and behavioral determinants* [Internal Report, Center "Leo Apostel"]. Free University of Brussels. <https://www.researchgate.net/publication/2420048>
- Howard, R. M. (2014). *Writing matters: A handbook for writing and research* (2nd ed.). McGraw Hill.

- Hyland, K., & Jiang, F. K. (2017). Is academic writing becoming more informal? *English for Specific Purposes*, 45, 40–51. <https://doi.org/10.1016/j.esp.2016.09.001>
- Jitpraneechai, N. (2019). Noun phrase complexity in academic writing: A comparison of argumentative English essays written by Thai and native English university students. *LEARN Journal: Language Education and Acquisition Research Network*, 12(1), 71–88. <https://so04.tci-thaijo.org/index.php/LEARN/article/view/168582>
- Karlsson, S. (2008). *Formality in websites: Differences regarding country of origin and market sector* [Master's thesis, Högskolan i Skövde]. Digitala Vetenskapliga Arkivet. <https://www.diva-portal.org/smash/get/diva2:37517/FULLTEXT01.pdf>
- Keppens, M. (2015). *A comparative study of formality in academic papers, conference papers, TED talks and popular science texts* [Master's thesis, Ghent University]. Ghent University Library. [https://libstore.ugent.be/fulltxt/RUG01/002/212/532/RUG01-002212532\\_2015\\_0001\\_AC.pdf](https://libstore.ugent.be/fulltxt/RUG01/002/212/532/RUG01-002212532_2015_0001_AC.pdf)
- Liardét, C. L., Black, S., & Bardetta, V. S. (2019). Defining formality: Adapting to the abstract demands of academic discourse. *Journal of English for Academic Purposes*, 38, 146–158. <https://doi.org/10.1016/j.jeap.2019.02.007>
- Nowson, S., Oberlander, J., & Gill, A. J. (2005). *Weblogs, genres, and individual differences* [Doctoral thesis, University of Edinburgh]. Edinburgh Research Archive. <https://era.ed.ac.uk/bitstream/handle/1842/1113/thesis.pdf?sequence=1&isAllowed=y>
- Praminatih, G. A., Kwary, D. A., & Ardaniah, V. (2018). Is EFL students' academic writing becoming more informal? *Journal of World Languages*, 5(2), 101–112. <https://doi.org/10.1080/21698252.2019.1570664>
- Prapobratanakul, C. (2024). Thai EFL undergraduate Engineering students' perspectives on academic writing: Challenges and strategies. *PASAA Journal*, 68, 1–31. <https://doi.org/10.58837/chula.pasaa.68.1.1>
- Schmid, H. (1994). Part-of-speech tagging with neural networks. *Proceedings of COLING 1994: The 15th International Conference on Computational Linguistics*, 1, 172–176. <https://aclanthology.org/C94-1027.pdf>
- Swales, J. M., & Feak, C. B. (2012). *Academic writing for graduate students: A course for non-native speakers of English* (3rd ed.). University of Michigan Press.
- Tocalo, A. W. I., Racman, S. P., Guiamelon, K. N. A. G., & Mama, B. M. (2022). Is Filipino ESL academic writing becoming more informal? *PASAA Journal*, 64, 99–125. <https://doi.org/10.58837/chula.pasaa.64.1.5>

## Appendix

### Examples of Formal and Informal Features in Students' Writing

Formal Features	
Noun	<ul style="list-style-type: none"> <li>- a key role in storing memories3. Without enough <u>sleep</u>, a person can have a hard time rememberi 1st_2017.txt</li> <li>- Here are the steps to get over a <u>breakup</u>. First, you would be in the Denial 1st_2019.txt</li> <li>- <u>people</u> are enjoying the movie together, and those <u>people</u> actually improve your experience. Some film genre 1st_2020.txt</li> </ul>
Adjective	<ul style="list-style-type: none"> <li>- community, or the company? If the change is <u>negative</u>, are there any other choice? To answer these 1st_2018.txt</li> <li>- to people around them less and not pay <u>enough</u> attention to their surroundings. Instead of commu 1st_2019.txt</li> <li>- share videos with the world. Not even a <u>single</u> audition or selection is required. The s 1st_2021.txt</li> </ul>
Informality Features	
Verb	<ul style="list-style-type: none"> <li>- everything for them, they can <u>decide</u> independently and fear of making their parents di 1st_2017.txt</li> <li>- In fact, it can <u>cause</u> a larger pore size which is a cause 1st_2019.txt</li> <li>- The massive stray dog packs are likely to <u>show</u> aggressiveness and be more encouraged to bite mor 1st_2021.txt</li> </ul>
Adverb	<ul style="list-style-type: none"> <li>- of education but not for teaching English. <u>Consequently</u>, the students didn't have a chance to 1st_2018.txt</li> <li>- you can be sure that the product is <u>definitely</u> available. On the contrary, when I was sho 1st_2020.txt</li> <li>- d makes children lose their identity. They cannot <u>truly</u> express their own needs, which makes kids have 1st_2021.txt</li> </ul>