



Informality in Applied Linguistics Research Articles: Comparing Native and Non-Native Writings

Mohammad Alipour ^{a*} , Mona Nooreddinmoosa ^{a †} 

^a *Islamic Azad University, Department of English Language Teaching, Ahvaz Branch, Ahvaz, Iran*

Received 28 May 2018 | Received in revised form 28 June 2018 | Accepted 6 August 2018

APA Citation:

Alipour, M., & Nooreddinmoosa, M. (2018). Informality in applied linguistics research articles: comparing native and non-native writings. *Eurasian Journal of Applied Linguistics*, 4(2), 349-373. doi: 10.32601/ejal.464196

Abstract

This quantitative-qualitative study aimed to fathom out whether and how informal features are exploited in articles of applied linguistics written in English by natives and non-natives. To this end, a corpus of 200 articles was compiled. We employed the classification of informal features proposed by Chang and Swales (1999) representing 10 informal features in academic writing. The AntConc software was used, along with manual search, to detect the informal features. The frequency, percentages, and the density per 1000 words of each informal feature were calculated. The results revealed that informal features are utilized more frequently in native articles than non-native ones, with no significant differences in the two corpora in terms of their most and least frequent informal features. Sentence *initial conjunctions* are the most recurrent informal features, while *exclamation marks* are employed the least frequently in both native and non-native articles. Implications for EAP courses are delineated in the study as well.

Keywords: Informal features; applied linguistics; research articles; natives; non-natives

1. Introduction

There appears to exist in a wide range of domains where language is applied a propensity for an alteration, though subtle, from formality to informality, conspicuous in online interactions, news coverage, business reports, and even presidential debates. This swing emboldens us to cast off decrepit orthodoxies and embrace more laid-back linguistic practices. Along the same line, so does academic writing seem to have been swept by this trend as research articles of various disciplines have been progressively becoming less formal (Hyland & Jiang, 2017), an area conventionally characterized by a remarkable degree of formality in writing. This used-to-be ubiquitous but now-perishing strong disposition toward formality is manifested by Coffin, Curry, Goodman, Hewings, Lillis, and Swann (2003) who called for the deployment of technical and abstract words and complex sentences in the majority of academic prose.

Given that there is an increased alacrity on the side of non-native English speakers for English for Academic Purposes (EAP), it is indispensable to execute studies to dig out the linguistic features that mold academic texts (Martinez, 2005). The impetus of

* Corresponding author. Tel.: +98-916-305-3776

E-mail address: alipour83@yahoo.com

† mona.nooreddinmoosa@gmail.com

the current study, hence, is to throw some light on and probe into the nature of this fickle notion, namely informality, in research articles of applied linguistics by exploring two corpora of articles developed by English native and Iranian non-native academics. We make attempts to unravel whether and to what extent they might resort to informality features, such as first person pronouns ‘*I*’ or ‘*we*’, to craft their writings and, if so, whether there are significant differences between these two groups in their employment of informality. In a nutshell, we seek to build a milieu to scrutinize informality in academic texts and grasp a better understanding of this concept in research articles. We believe that this can be accomplished by measuring the occurrence of certain informal features and then juxtaposing non-native writers’ articles with those by native writers as the benchmark data. To fulfill these ends, we execute a quantitative analysis of how informality behaves in research articles of applied linguistics written in English and a qualitative study to compare it between native British English writers and Iranian non-native writers. Further, we seek to link our findings to those of Hyland and Jiang (2017) in order to check the appropriate use, underuse, or overuse of specific informality features through text analysis statistics, particularly occurrence densities and log-likelihood indices.

This paper proceeds with the literature review to delineate informality and sketch some studies conducted in this realm, coupled with the objectives of the study. We move on to introduce the method adopted to execute this research, followed by the results and the pertinent statistics and tables. It culminates in a discussion of the findings and a set of conclusions drawn accordingly.

2. Review of literature

2.1. *Informality versus formality*

It is assumed that writing is universally and increasingly losing its formality in many realms (e.g. Adel, 2008; Fairclough, 2001; Foster, 2005). Mair (1998, p. 153) pointed out to “a trend towards the informal and the colloquial in written communication” and Leedham (2015) found more prominent informality in undergraduate essays. Fairclough (2001, p. 52) proposes “synthetic personalization” means this tendency to informality could be thought of as an aspect of a modern idea which clouds explicit hierarchies and cherishes being interpersonally involved; on the other hand, it can be seen as another form of stealthy persuasion (Hyland & Jiang, 2017). Atkinson (1999) attempted to examine informality by executing a longitudinal corpus-based study to detect changes in articles covering three centuries between 1675 and 1975. Their findings revealed that informal features soared unremittingly, implying a shift in rhetoric from author-centeredness to object-centeredness.

Academic writing is believed to possess an inclination not to breach the formal writing etiquette of disinterest and fastidiousness, enabling writers and readers to exude an air of detachment. This is accomplished when authors strive to strike a sense of anonymity and egalitarianism in their textual products with recourse to formality features in their quest for veracity and objectivity. Hyland and Anan (2006), for instance, compared native and non-native teachers’ attitudes toward formality in English writing and uncovered that Japanese ones show a stronger tendency to

consider stylistic deviations as erroneous, while native tutors were more sensitive to formality, reproaching informality as academic inappropriacy. Notwithstanding the belief that portrays informality as an aberration from this academic impartiality, it should not be regarded as opposing formality, rather as the instantiation of a reader-friendly approach to text generation and an opening for the infusion of a more subjective tenor. Nonetheless, informality, which has permeated a large number of oral and written fields, can be sometimes described in light of formality and has expanded to writing academic texts (Hyland & Jiang, 2017).

Heylighen and Dewaele (1999, p. 1) believe that, “a formal style is characterized by detachment, accuracy, rigidity and heaviness; an informal style is more flexible, direct, implicit, and involved, but less informative”. They also mentioned that the formality of academic prose contributes to keeping away equivocation and misapprehension by decreasing the dependence on context and ambiguity of definition, whereas its informality dismisses pompous conventionality to build a comfortable and friendly rapport.

Informality assumes that formality exists and that there is an accepted constellation of routines which rest upon a system, construction, or specialist (Hyland & Jiang, 2017). Therefore, formal language is delineated in terms of being “very correct and serious rather than relaxed and friendly” whereas it deals with ‘negative politeness’ in pragmatics and the deployment of disaffecting behavior to esteem the face of another side and circumvent imposition on the face their own (Brown & Levinson, 1978).

In accord with Heylighen and Dewaele (1999), academic texts lie on a cline in terms of formality/informality and circumvention of vagueness. From their vantage point, informal language deals with the necessary background knowledge for the realization of communication; however, complete informal fuzziness “merely signifies that any interpretation is as likely as any other one” (Heylighen & Dewaele, 1999, p. 9). They hold that formality in academic writing is pivotal to intelligibility on the ground that it leaves no wiggle room for ambiguity caused by the author’s personal quirks. Informality, which is basically taken as “the absence of full grammatical sentences, a decrease in concern about punctuation, and a high tolerance for typographic and spelling errors” (Coffin et al., 2003, p. 141), should not be looked upon as the relinquishment of these writing conventions, but as the writers’ endeavor to create a congenial atmosphere wherein they can breathe a sigh of a collegial familiarity with readers and unleash themselves from the shackles of academic writing bigotry.

Tenor, or the selection of grammar, also has to do with formality, and this allows language users to achieve their intricate and different interpersonal relationships when they opt for language choices which manifest a suitable facade and a proper affinity with readers (Halliday, 1985). Hence, this deals with constructs like colloquialism (e.g. Hundt & Mair, 1999) or the use of language by average users to indulge in day-to-day interactions, and engagement, or the way via which writers endeavor to connect to the readers (Hyland, 2005). One should not, nevertheless, surmise that the gap between academic writing and conversation is being bridged (Hyland, 2004). In academic texts, to determine informality, it is conceivable to think

of a number of elements which are high likely to co-occur, despite the fact that there is no general consensus on the nature of these features (Hyland & Jiang, 2017).

Nowadays a wide gamut of written and oral areas of discourse is infused with informality, with academic texts pursuing this tendency (Hyland & Jiang, 2017). In accord with Hyland (2004), applied linguists increasingly tend to seek proof for further interactivity in academic writing to recognize how researchers forge a comprehensive alliance with readers. Studies have reported a steady shift of pendulum from typical inimical and detached writing styles to those which permit stylistic variation and a more intimate remark and account. Mair (1998, p. 153), for instance, reports “a trend towards the informal and the colloquial in written communication”. Leedham (2015) discovered more informality in the writings of undergraduate students. Academic genres, nonetheless, seem to be fairly impervious to diffusion by colloquial elements (Seone & Loureiro-Porto, 2005).

Broadly speaking, the ingredients of informality vary from those of formality. A vivid illustration of such a disparity is the study by Gilbert and Mulkay (1984) suggesting that scholars resort to written and oral means to unveil their trains of thoughts. For one, they discover scholars are inclined to beset their interviews with polemics, analytical acumens, and societal bigotries. Authors refer to this as a dependent collection which excludes impartial reactions to the accepted environment, but entails the judgments of individuals who behave in accordance to their specific social circumstances. By way of contrast, the same writers’ research products are exempt from these features, but tend to be ruled by an additionally neutral empiricist repertoire: “Empiricist discourse is organized in a manner which denies its character as an interpretive product and which denies that its author’s actions are relevant to its content” (Gilbert & Mulkay, 1984, p. 56).

One can resort to Biber’s (1988) multidimensional corpus analysis to detect connected features in spoken and written texts. Employing factor analysis, he illustrates the ways in which 16 main types of grammatical normally go together in five proportions of variation, with the first dimension, called “involved in compared to informational production”, closest to the notion of informality/formality. The categories included greater interjections and adverbs pronoun and verb forms, and fewer prepositions, nouns and attributive adjectives. He suggest these “reflect direct interaction, give attention to instant circumstance and personal attitudes or feelings, fragmentation or decrease in form, and a less specific, generalized context” (1995, p. 143). Although he explains the divergences as spoken versus literate text, they typify a typical discourse which evocates bosom interactions, doing away with the compulsion of a shared context. These contribute to streamlining the notion of formality and how it is manifested (Hyland & Jiang, 2017).

Another approach to formality has found it in building up the facets which help a text to be scrupulous and liberated from its context. Therefore, they put forth a recipe to evaluate formality through making an inventory of total words with an ostensible purpose in relation to the communicative context in one group and aspects which are commonly manifest of understanding of such a setting. It is claimed that when nouns, adjectives, prepositions and articles recur with greater frequencies, this diminishes

the amount of contextual information required to apprehend a text and thus augments the formality of the text; nevertheless, more verbs, pronouns and interjections intensify the necessity of the immediate context and curtail the formality of the text, signifying that additional context provision results in higher text formality. This goes a long way to show that it is infeasible to produce an entirely formal text which cannot establish a context and build a rapport with readers. (Heylighen & Dewaele, 1999).

Texts can be considered to fall along a continuum with two extremes of complete obscurity and complete scrupulousness where the writer's context appraisal yields influence on the accurate degree of formality. As a consequence, Heylighen and Dewaele (1999) follow Biber (1995) to characterize formal texts in terms of incorporating greater nominalization and typify informal ones with respect to greater verbalization. Provided that this method holds true, formality is regarded synonymous with liberation from context, and it is held to be attainable via the function of grammatical classes. However, despite the fact that pronouns can be applied to deixically link an utterance or a proposition to a particular context, and are thus construed as fitting the 'informal' extreme, one needs to take heed in fathoming that verbs are less formal than nouns. As a matter of fact, noun complement structures can inject more informality into texts as they bring to the fore a writer's position toward a message (Jiang & Hyland, 2015).

Informality features can be pinpointed through making use of style guides, which echo the apprehensions and attitudes of teachers, students and experts (Hyland & Jiang, 2017). Benet (2009) studied style manuals and revealed that the formal personality of academic writing is accomplished via extensive use of Latinate vocabulary and detached forms, with numerous mentions of 'objectivity'. In another study, Chang and Swales (1999) examined style booklets which aimed at developing an inventory of grammatical elements to fulfill an adequate amount of formality. The list of 'informal elements' embraced WH questions, listing expressions (such as etc., and so on), and first person pronouns.

2.2. Objectives of the study

In line with these studies, particularly the one by Hyland and Jiang (2017), it can be assumed that applied linguistics articles like those of a range of academic fields have become more informal over the past years. Since formality/informality can be seen along a continuum, repudiating polar extremes, it is gratifying to count the number of features and rhetorical characteristics to size up informality in academic writing. Accordingly, the question to put to test is whether non-native authors are capable of following native authors in this recent trend to appropriately employ informality in their articles. This can be exposed by comparing articles developed by natives and non-natives, which is the motivation of the current study. In this study, among the many genres of academic language, we opt for research articles since they are regarded as the most salient academic writing genre by many academics (e.g., Belcher, 2007; Hyland, 2005; Hyland & Jiang, 2017; Kahkesh & Alipour, 2017; Moreno, 2010)

In recent years, there has been an exponential upsurge in the pervasiveness of English as a lingua franca whose popularity has increasingly grown in English as a Foreign Language (EFL) contexts (Aufa, 2014), such as Iran. The significance of English in academia is undeniable with general consensus over it among academics (Belcher, 2007; Hyland, 2009; Jenkins, 2011; Lillis & Curry, 2010), contending that it is “unquestionably the language of international scholarship” (Hyland, 2009, p. 83). Albeit there have been calls for a divorce from native-speaker norms, echoed by scholars including Jenkins (2000, 2003) who champion this separation and a “time of shifting sands” (Jenkins, 2000, p. 49), we can witness that native-speaker standards are still preferable and idyllic, domineering these environments (Celik, 2006; Florence, 2012; Ozturk & Atay, 2010; Sifakis, 2014). Attempts for this paradigm shift can also be seen in academic writing despite the domination of native-speaker conventions in “the academic publishing world”, giving rise to “the marginalization of researchers who are not first-language speakers of English” (Strauss, 2017, p. 1).

Given the still-pressing stranglehold on academic writing by native writers which does not appear to be loosening up any time soon, Jenkins (2011) asserts that “the attachment to native English pervades global academic life” (p. 927). Strauss (2017), hence, holds the belief that it is futile to battle the tide and that non-native researchers should pragmatically concur with native-writer norms as they seem to form the most practical and well-worn path to take so as to publish in prestigious English-medium journals. In making her case for Spanish writers to publish in English journals, Moreno (2010) believes that “protesting against and criticizing mainstream practices would be a disservice to the scholars” (p. 58). She maintains that scholars need to abide by these native-writer regulations to be able to disseminate their articles in well-established journals of their fields which can aid them to gain international renown and professional promotion. Since Iranian researchers are also second-language speakers of English and are under the influence of the publish-or-perish mind-set of the academic world to “perpetuate and legitimize the infrastructures of power” (Mauranen, 2009, p. 215), it is incumbent upon us to undertake this study to develop a better understanding of how the native-writer standards of utilizing informality in academic writing contribute to text production.

Iranian authors are generally inclined to write articles in Persian very formally, circumventing self-mentions and personal pronouns (*I* and *we*) (Faghih & Rahimpour, 2009); instead, they lean towards applying third person pronouns (*he/she* and *they*), circuitously referring to the author and not his or her name (e.g. *the researcher believes that...*), or heavily utilizing passive structures (e.g. *a reading test was used as the pretest...*). This marginalization of the authors goes a long way to manifest the researcher’s unpretentiousness which is fortified by the Iranian culture. On the ground that there are some cultural differences between native and Iranian non-native writers, these cultural discrepancies may impose adverse effects on the ways Iranians write their articles in English when they endeavor to inject a high degree of formality in their texts and negatively smuggle formal features into English from Persian. This is of great substance since native speakers’ language and academic discourse conventions serve as the yardstick based on which the appropriateness of non-native writers’ academic texts including articles is evaluated.

To date, a number of studies have attempted to compare native and Iranian non-native articles written in English, especially in terms of metadiscourse markers (e.g., Kahkesh & Athour, 2017; Keshavarz & Kheirich, 2011; Mirzapour & Rasekh Mahand, 2012; Yeganeh, Heravi, & Sawari, 2015). Nevertheless, it appears that no serious research study has been executed in this context in order to juxtapose informality in applied linguistics articles written in English by native and Iranian non-native researchers. In this light and in accord with Hyland and Jiang's (2017) study as a springboard for ours, the purpose of this research undertaking is twofold. We first endeavor to find out whether Iranian writers can fittingly follow native writers' norms sketched above in using informal features with respect to the density with which they deploy them in applied linguistics articles in English. In other words, we try to tease out whether Iranians overuse, underuse, or make a judicious use of informality. We also make an attempt to see which types of informal features are more frequently used in the articles by each group of researchers, aspiring to cast light on the areas where non-native authors conform to and deviate from the norms of their native counterparts. The questions which shape the premise upon which this study is constructed are as follows: Are native writers inclined to employ some informality features more intensely and the others less? How about non-native writers? Can they appropriately conform to native speakers' writing norms in terms of informality? Do they overuse or underuse informality in general or some of its specific features in particular?

3. Methodology

3.1. Corpus

We compiled a corpus of 100 articles written by native English writers and also 100 articles by Iranian non-native writers in English. To control for cultural influences on the use of informality features, articles developed by merely British native speakers of English were incorporated in the corpus, leaving out those by native speakers of American English, Australian English, or natives from any other Commonwealth country. The Iranian authors were all native speakers of Persian who speak English as a foreign language. Writers' names, affiliations, and biographies were checked in each article and some emails were sent to writers in order to make sure all the articles were authored by British natives. All the articles were compiled from six journals of the applied linguistics discipline between 2012 and 2017. The compilation method of the native corpus was informed by judgment sampling described above from the large pool of articles in this time span, whereas for the non-native corpus we opted for simple random sampling. The native articles were collected from three well-established international journals, namely *Annual Review of Applied Linguistics* (IF=4.88), *Applied Linguistics* (IF=3.225), and *Applied Linguistics Review* (IF=1.286), all of which are ISI indexed journals with Impact Factor (IF). Likewise, the non-native corpus was composed of three accessible Iranian national journals, i.e., *Iranian Journal of Applied Language Studies*, *Journal of Research in Applied Linguistics*, and *Iranian Journal of Applied Linguistics*. These are all peer-reviewed journals, confirmed by Iran's Academic Journals Commission and indexed in Islamic Science

Citation Database (ISC). All the articles were based on the IMRD model (Swales, 1990), which is a standard format in articles partitioning the paper into "Introduction, Method, Result, and Discussion" (Nwogu, 1997). Once the articles were gathered, word count was run in order to determine the size of the corpora. It was revealed that the total number of words in the non-native corpus was 692046, while there were 910740 words in the native corpus. This discrepancy between the sizes of the corpora, however, does not jeopardize the reliability of the findings in the study by virtue of the fact that the data were balanced by means of computing the density of informality features per 1000 words. This is a common practice in corpus-based studies which yields a trustworthy and standard index to compare the results of text analysis in case the corpora are rendered compatible in light of the number of articles and not tokens (Adel & Erman, 2012; Alipour & Matouri, 2017; Durrant & Schmitt, 2009; Hyland, 2005; Hyland & Jiang, 2017; Martinez, 2005).

3.2. Procedure

This study was a corpus-based empirical one (Biber, Conrad & Reppen, 1999), which entailed quantitative analysis of an assemblage of natural texts via a text-analysis software, coupled with qualitative interpretation of the data. The native corpus functioned as the reference corpus with which the non-native corpus was contrasted (Flowerdew, 2001; Scott, 2001). We employed the informal features identified by Chang and Swales (1999) represented below, which was also applied by Hyland and Jiang (2017) to conduct their study. This list is compatible to the broad understanding of informal language (Biber, 1988; Chafe, 1986b; Nash, 1986) and entails such factors as interactiveness, participation, and reader engagement (Hyland & Jiang, 2017).

Table 1. Informal Features and Examples

1. First person pronouns to refer to the author(s) (I and we)
e.g., "I will approach this issue in a roundabout way."
2. Unattended anaphoric pronouns (this, these, that, those, it) that can refer to antecedents of varying length
e.g., "This is his raw material."
3. Split infinitives – an infinitive that has an adverb between to and the verb stem
e.g., "The president proceeded to sharply admonish the reporters."
4. Sentence initial conjunctions or conjunctive adverbs
e.g., "And I will blame her if she fails in these ways."
5. Sentence final preposition
e.g., "A student should not be taught more than he can think about."
6. Listing expressions ('and so on', 'etc', 'and so forth' used when ending a list)
e.g., "These semiconductors can be used in robots, CD players, etc."
7. Second person pronouns/determiners to refer to the reader (you and your)
e.g., "Suppose you are sitting at a computer terminal which assigns you role R"
8. Contractions
e.g., "Export figures won't improve until the economy is stronger."
9. Direct questions
e.g., "What can be done to lower costs?"
10. Exclamations
e.g., "This is not the case!"

We hired a bottom-up analysis method where we first recognized the features representing informality, identified their types, and then compared the features across the native and non-native corpora. The detection method was a combination of electronic analysis and human intervention. Most of the informality features were spotted by the AntConc software (Anthony, 2011). To pin down each feature, a key word or a sign was fed into AntConc, which demonstrated each key word or each sign in the articles by highlighting it. Having come up with instances of informality through the software, we then had to read them all one by one and manually check them, as in Howarth (1998), against the framework to ascertain that they indeed stood in as an informality feature. To illustrate this, as for *direct questions* or *exclamations*, we keyed in a question mark (?) or an exclamation mark (!) in the search bar of the software and afterward discarded the ones which did not match the criteria. Concerning *contractions*, for example, we typed in the apostrophe sign (') and looked through all the instances yielded by the software, shedding those which did not fit in as a *contraction*.

It should be pointed out that, in order to locate the possible errors and to ensure the reliability of the data analysis method, we first piloted the study based on ten percent of the data. To do so, this part of the data was scrutinized separately by each researcher. Inter-rater reliability of the analysis was then calculated through Cohen's Kappa correlation, revealing a coefficient of ($k=0.87$) which confirmed the reliability of the analysis method.

Finally, the frequency of each element was counted and written, and after calculating the percentage of each informal feature, its density per 1000 words was also computed. The frequency and the density of each type of the informal features in each corpus were then compared.

4. Results

In this section, we present the results of the analysis in the following tables. In what follows we first compare the data of the non-native corpus vis-à-vis the native corpus through frequencies, percentages, densities and the chi-square test for association. Afterward, the non-native data are compared with the native data and those in Hyland and Jiang's (2017) study to locate the areas of overuse or underuse by non-native writers by way of log-likelihood and log-ratio.

4.1. Non-native articles versus native articles (test of significance)

The raw frequencies, percentages, and the densities (per 1000 words) of the informal features are summarized in the tables below.

Table 2. Details of the Corpora

	Non-Native articles	Native articles
Total number of words	692046	910740
Total number of informal features	6527	10090
Density per 1000 words	9.43	11.07

In the comparison between the native and non-native corpora, Table 2 indicates that out of 692046 words in the non-native articles, 6527 signaled informality. On the other hand, out of 910740 words in the native articles, 10090 informal features were detected. As illustrated by the log-likelihood index in Table 11 below, it can be seen that informal features were utilized more intensively by the native writers, indicating an underuse of these feature by the non-native writers.

Table 3. Results of Informal Features in Native and Non-Native Articles

Informal features	Non-Native Articles		Native Articles	
	Frequency (%)	Density	Frequency (%)	Density
1. First person pronouns	455(6.97)	0.65	1980(19.62)	2.17
2. Unattendant anaphoric pronouns	2022(30.97)	2.9	3346(33.16)	3.67
3. Split infinitives	23(0.35)	0.03	95(0.94)	0.10
4. Sentence initial conjunctions	3174(48.67)	4.58	3947(39.15)	4.33
5. Sentence final prepositions	49(0.75)	0.07	46(0.45)	0.05
6. Listing expressions	39(0.59)	0.05	159(1.58)	0.17
7. Second person pronouns	343(5.25)	0.49	141(1.39)	0.15
8. Contractions	46(0.7)	0.06	73(0.72)	0.08
9. Direct questions	350(5.36)	0.50	277(2.74)	0.30
10. Exclamations	26(0.39)	0.03	26(0.25)	0.02

According to Table 3, *sentence initial conjunctions* were the most frequent informal feature used by the writers of both native and non-native articles. They were used 3174 times, which equals 48.67 % of all the informal features in the Iranian non-native corpus, and 3947 times amounting to 39.1% in the native corpus. Nonetheless, *exclamation marks* were the least frequent informal feature in both corpora, employed 26 times (0.39%) in the Iranian non-native articles and 26 times (0.25 %) in the native articles.

In order to tease out whether or not the differences in the use of informality features across the two corpora were statistically significant, we carried out a significance test, chi-square test for independence ($\alpha < .05$), which in the present study illustrates the relationship, if any, between the use of informality in academic writing and being a native or non-native writer. It is worth noting that, due to the gap in the total number of words in each corpus, the data in the non-native corpus were normalized on the basis of the data in the native corpus prior to being fed into SPSS (version 21).

Table 4. Result of Chi-square

Person Chi Square for Native and Non-Native Articles			
Pearson Chi Square	Value 1013.412 ^a	Df 9	Asymp.sig. (2-sided) .000

As can be seen in Table 4, chi-square ($0.0 < .05$), which was conducted for all informality features, also manifests significant differences in the frequency rates of the informal features in the two corpora.

Table 5. The Standardized Residuals for Native and Non-Native articles

Informal features	Non-Native articles		Native articles		Total Count
	Count	Std.Residuals	Count	Std.Residuals	
1. First person pronouns	598	-17.1	1980	15.7	2578
2. Unattendant anaphoric pronouns	2660	-1.9	3346	1.8	6004
3. Split infinitives	30	-3.6	95	3.3	125
4. Sentence initial conjunctions	4177	7.2	3947	-6.7	8124
5. Sentence final prepositions	64	1.9	46	-1.7	110
6. Listing expressions	51	-4.6	159	4.3	210
7. Second person pronouns	451	10.8	141	-10.00	592
8. Contractions	60	-0.1	73	0.1	133
9. Direct questions	460	6.6	277	-6.1	737
10. Exclamations	34	1.2	26	-1.1	60

As Table 5 reveals, the standardized residuals for the *sentence final proposition*, *contractions*, *un attendant anaphoric pronouns*, *exclamation marks* are between + 1.96 and – 1.96, illustrating no significant differences between non-native and native articles in these features. However, there seems to be significant differences in the use of the other informal features between the two corpora, by having the standard residuals which are more than +1.96 and less than -1.96. Therefore, they can be said that these features account for the differences in the use of informal features between native and non-native articles as depicted by the chi-square above.

4.2. Log-likelihood and effect size

Log-likelihood is a measure of statistical significance which allows the researcher to decide if a difference exists between two corpora, but not revealing how large or how significant the difference is. This is a sore issue which can be ameliorated by means of the Log Ratio which is an effect-size statistic, demonstrating the magnitude of the variance between two corpora for a given feature by comparing its relative frequency in one corpus to that of the same feature in a second corpus. In simple terms, Log Ratio means how many times bigger the relative frequency is in one corpus relative to the other (Hardie, 2014; Johnston, Berry, Mielke, 2006; Reyson, 2008). This section embraces a blow-by-blow analysis of overuse or underuse of the above informality features in the non-native corpus relative to two other corpora, namely the native corpus compiled in the present research and the one in Hyland and Jiang's (2017) study. Hyland and Jiang's corpus of applied linguistics research articles in English included 30 articles published in 2015 with a total number of 237,452 words. By and large, the results which cropped up in our investigation were not symptomatic of any appropriate use of the features by the non-native writers.

We bring up each informality feature below and provide the figures pertaining to the log-likelihood and the Log Ratio. It should be noted that, in the following tables, (%1) and (%2) values depict relative frequencies in the non-native corpus and the native corpus/Hyland and Jiang's corpus respectively. Moreover, (+) indicates overuse in the non-native corpus and (-) underuse. Furthermore, if a feature has a comparable relative frequency in the no-native corpus and the native corpus/Hyland and Jiang's

corpus, the binary log of the ratio is 0. In case it is 2 times more frequent, the binary log of the ratio equals 1, and if it is 4 times more recurrent, the binary log of the ratio amounts to 2, and so forth.

4.2.1. Cases of overuse

Of the ten features scrutinized for log-likelihood, *initial conjunctions*, *second person pronouns*, *sentence final prepositions*, *exclamations*, and *direct questions* were overused by the Iranian writers compared to the native British English writers, who make up the main touchstone corpus. Nevertheless, apart from *initial conjunctions* and *direct questions*, they all showed underuse relative to Hyland and Jiang's (2017) corpus.

Table 6. Log-likelihood and Effect Size for Sentence Initial Conjunctions

	%1	%2	LogRatio		%1	%2	LogRatio
Non-native	0.46		0.08	Non-native	0.46		1.21
Native		0.43 +		Hyland & Jiang	0.38 +		

Table 7. Log-likelihood and Effect Size for Second Person Pronouns

	%1	%2	LogRatio		%1	%2	LogRatio
Non-native	0.05		1.68	Non-native	0.05		-1.01
Native		0.02+		Hyland & Jiang	0.10 -		

Table 8. Log-likelihood and Effect Size for Sentence Final Prepositions

	%1	%2	LogRatio		%1	%2	LogRatio
Non-native	0.01		0.49	Non-native	0.01		-0.45
Native		0.01+		Hyland & Jiang	0.01 -		

Table 9. Log-likelihood and Effect Size for Exclamations

	%1	%2	LogRatio		%1	%2	LogRatio
Non-native	0.9		0.40	Non-native	0.9		-1.09
Native		0.00 +		Hyland & Jiang	0.01 -		

Table 10. Log-likelihood and Effect Size for Direct Questions

	%1	%2	LogRatio		%1	%2	LogRatio
Non-native	0.05		0.73	Non-native	0.05		4.59
Native		0.03 +		Hyland & Jiang	0.0 +		

As evident in the tables, the most conspicuous cases of informality overuse by the non-natives were *second person pronouns* and *direct questions* based on the Log Ratios

which indicate that they were made use of by Iranian writers approximately three times and twice more respectively in comparison to the natives, whereas the ratio pertaining to *initial conjunctions* (0.08) is manifest of a negligible overuse, which can also be regarded as the only case of appropriate use of informality. With respect to Hyland and Jiang’s (2017) corpus, *initial conjunctions* and *direct questions* were also applied with markedly greater frequencies by the non-natives, while they underused the other features.

4.2.2. Cases of underuse

The results divulge that the overall occurrence rate of informality features in the non-native articles was an instance of underuse of roughly 25 percent compared to the native corpus and 100 percent relative to Hyland and Jiang’s (2017) corpus according to the effect sizes in Table 11. The features which represent underuse by the Iranian authors vis-à-vis the native and Hyland and Jiang’s corpora include *first person pronouns*, *unattended anaphoric pronouns*, *listing expressions*, *contractions*, and *split infinitives*.

Table 11. Log-likelihood and Effect Size for Overall Informal Features

	%1	%2	LogRatio		%1	%2	LogRatio
Non-native	0.94		-0.23	Non-native	0.94		-1.02
Native		1.11 -		Hyland & Jiang		1.91 -	

Table 12. Log-likelihood and Effect Size for First Person Pronouns

	%1	%2	LogRatio		%1	%2	LogRatio
Non-native	0.07		-1.73	Non-native	0.07		-2.99
Native		0.22 -		Hyland & Jiang		0.52 -	

Table 13. Log-likelihood and Effect Size for Unattended Anaphoric Pronouns

	%1	%2	LogRatio		%1	%2	LogRatio
Non-native	0.29		-0.33	Non-native	0.29		-1.29
Native		0.37 -		Hyland & Jiang		0.72 -	

Table 14. Log-likelihood and Effect Size for Listing Expressions

	%1	%2	LogRatio		%1	%2	LogRatio
Non-native	0.01		-1.63	Non-native	0.01		-1.75
Native		0.02 -		Hyland & Jiang		0.02 -	

Table 15. Log-likelihood and Effect Size for Contractions

	%1	%2	LogRatio		%1	%2	LogRatio
Non-native	0.07		-0.27	Non-native	0.07		-1.02
Native		0.01 -		Hyland & Jiang		0.13 -	

Table 16. Log-likelihood and Effect Size for Split Infinitives

	%1	%2	LogRatio		%1	%2	LogRatio
Non-native	0.0		-1.65	Non-native	0.0		-2.66
Native		0.01		Hyland & Jiang	0.02	-	

In accord with the figures in Tables 12 to 16, comparing to the native articles, *first person pronouns*, *listing expressions*, and *split infinitives* are the feature which were most substantially underused in the non-native articles with nearly three to four times lower deployment rates (Log Ratios=1.73, 1.63, and 1.65 respectively). The same also holds water for Hyland and Jiang's (2017) corpus where all these features were employed more frequently in their corpus than ours, somewhere in the vicinity of twice (*contractions*), three times (*unattended anaphoric pronouns* and *listing expressions*), five times (*split infinitives*), and eight times (*first person pronouns*) more.

4.3. How does informality behave?

Overall, the results indicate that concerning the different use of informal features in the native and Iranian non-native articles, authors employ informal features such as *first person pronouns*, *listing expressions*, *second person pronoun*, *split infinitives*, *direct questions* and *sentence initial conjunctions* with different frequencies. Therefore, while informal features are employed more in applied linguistics research articles written by native writers, this is owing to some informal features and hangs on the two corpora and the features in focus. Even in our corpus, the differences in using informal features do not transpire in all features and there are no significant differences between non-natives and natives in employing *sentence final proposition*, *contractions*, *unattendant anaphoric pronouns*, *exclamation marks*.

As Table 3 portrays, *first person pronouns*, *unattendant anaphoric pronouns*, *sentence initial conjunctions*, *second person pronouns*, and *direct questions* comprise the major bulk of informality features in both corpora, accounting for 97.22 percent of all elements in the non-native corpus and 96.06 percent in the native corpus. That is why we merely discuss each of these features and dispense with the discussion of the other features as they altogether make up a minute proportion of informality combed for in the articles.

First person pronouns: They are pronouns which refer to the name(s) of author(s) (Chang & Swales, 1999). Personal pronouns are primarily exploited in research articles to illuminate the work developed by the author and set out the research procedure (Hyland, 2001, p. 220). These are strategic assets which are the key to building the writer's facade by placing them in their academic discipline and contributing to their image of expertise (Kuo, 1999; Tang & John, 1999). They are the second most frequent element in Hyland & Jiang's (2017) corpus, but the third most recurrent one in our non-native and native corpora. The striking point about them, however, is that Iranian English writers tend to underutilize them in articles of applied linguistics comparing to native British writers by approximately one third. This underuse is even more discernible relative to Hyland & Jiang's corpus by roughly one sixth. Table 17 manifests the use of *I* and *we* in our corpora, showing that Iranian non-natives generally fail to comply with native speakers' norms of applying these

pronouns both in terms of the rate of using them and also the type of pronoun, as they make use of very few cases of *I*.

Table 17. Details of First Person Pronouns in the Non-native and Native Corpora

	Non-native corpus		Native corpus	
	Frequency	Density	Frequency	Density
I	59	0.08	686	0.75
We	396	0.57	1294	1.42

This disposition observed in the present study is in line with previous studies on the use of *first person pronouns* by non-native speakers from different backgrounds, for example, Hyland's (2002a) study of non-native undergraduate writings, Ivanic and Camps's (2001) study of Mexican postgraduate students' academic writings, and Martinez's (2000) study about writings by Spanish non-native speakers of English. Attarn (2014) and Zarei and Mansoori (2011) also presented that self-mentions are applied more by native writers than non-native writers. Furthermore, Martinez (2005) holds that this resource is highly challenging for non-native English-speaking writers.

Our observation runs counter to Hyland and Jiang (2017) who found an increase in the use of these pronouns, claiming that it stems from the more pronounced participation of non-native speakers of English in publishing articles on the ground that they are not as sensitive as native speakers to personal projection and authority denoted by *first person pronouns*. Hyland and Jiang's discovery is also backed up by Seone (2013) who traced that the surge in the use of *I* and *we* is associated with a plunge in the use of *passives*. It appears that our outcome conforms to Seone because English articles written by Iranians are generally characterized by a great degree of passivation although we did not scrutinize passives in our analysis, but through personal encounters and experience.

It can be noticed that the reasonable explanation for this difference between native and non-native writers can stem from cultural differences. Iranian academics generally believe that resorting to *first person pronouns* is an inappropriate strategy based on their culture and prefer to use *third person pronouns*, expressions such as *the researcher(s)*, or *passive structures* to introduce themselves. On the other hand, native writers tend to be more direct than indirect, so they show a stronger inclination to make use of these features. Some researchers, such as Kaplan (1966) and Hofstede (1991) asserted that Iranians tend to be more indirect than direct. In the examples below, the pronoun *I* refers to the current writer and the pronoun *we* refers to more than one writer.

Ex 1: **We** can gain a more nuanced view of the spread of vocabulary across sub-corpora by quantifying the overlaps between each. (Native corpus, Durrant, 2014)

Ex 2: You will recall from the introduction section that **I** regarded Iran and Saudi Arabia's changes in areas of focus, followed by changing their discursive tones, as dialectical variation. (Non-native corpus, Vakili Latif, 2016)

Sentence initial conjunctions: These elements are inclined to transpire mostly in spoken language and spontaneous language production (Chafe, 1986b). Concerning the frequencies and occurrences of informal features, *sentence initial conjunctions* are the most frequent type of informal features in both native and non-native articles and are overemployed by non-native writers, though slightly. This overuse is also perceived in the writings of Chinese (Milton & Tsang, 1993), Cantonese (Field & Yip, 1992), and Norwegian (Evensen & Rygh, 1988) non-native speakers of English. Based on Tables 18 and 19, natives and non-natives also differ in terms of instances of conjunction although the most recurrent element in both corpora is *however*, with a density greater than 0.47. It is also the element with the highest density in Hyland and Jiang's (2017) corpus. Of the most frequent conjunctions in the native articles, *for example*, *indeed*, *hence*, and *so* are absent in the non-native articles, and instead Iranian writes show a tendency to resort to *therefore*, *furthermore*, *moreover*, and *thus*.

Table 18. Most frequent Initial Conjunctions and Conjunctive Adverbs in the Non-native Corpus

Non-native corpus	Frequency	Density
However	323	0.46
Therefore	270	0.39
Furthermore	171	0.24
Moreover	99	0.14
Also	92	0.13
Finally	85	0.12
First	85	0.12
In addition	79	0.11
Thus	66	0.09
In fact	66	0.09

Table 19. Most frequent Initial Conjunctions and Conjunctive Adverbs in the Native Corpus

Native	Frequency	Density
However	435	0.47
For example	257	0.28
Indeed	112	0.12
First	112	0.12
Also	66	0.07
Hence	52	0.05
Finally	46	0.05
So	39	0.04
In addition	33	0.03
And	19	0.02

The difference in using this feature in the two corpora may refer to the identities of these differing groups of writers. Hyland stresses that successful academic writing does not occur in a vacuum but hinges on a writer's estimation of a context. This is not to say that these conventions are fixed and monolithic – they change over time in response to changing circumstances – but discursive innovations emerge slowly and

disciplinary conventions are a valued symbolic resource in identity construction (Hyland, 2009). The following examples show how conjunctions are applied.

Ex 3: **Moreover**, regression analysis revealed that significant difference exists between EFL learners' self-efficacy and metacognitive awareness in predicting use of language learning strategies in a way that metacognitive awareness entered the model as the best predictor of language learning strategies. (Non-native corpus, Ghafournia & Afghari, 2013)

Ex 4: **However**, it is not the aim of conducting experiments in CDA to claim that the social, political, and historical dimensions of discourse are not important in the way that texts are interpreted. (Native corpus, Duff, 2015)

Unattended anaphoric pronouns: These pronouns are also basically attached to spoken discourse and writers are advised to abstain from them (Geisler, Kaufer & Steinberg, 1985). Nonetheless, Swales and Feak (2012) argue that *anaphoric pronouns* enable writers to sound more dexterous and imposing. They are also found to occur frequently in academic texts thanks to their flexibility to refer to different segments of a text (Biber, et al., 1999) since these backward-pointing pronouns can refer to antecedents of any length (Hyland & Jiang, 2017). Iranian non-native writers of English appear to mildly underuse these pronouns in comparison to native writers and Hyland and Jiang's (2017) corpus.

Table 20. Most frequent Unattended Anaphoric Pronouns in the Non-native and Native Corpora

	Non-native corpus		Native corpus	
	Frequency	Density	Frequency	Density
This	376	0.54	990	1.09
That	250	0.36	360	0.39
These	76	0.19	205	0.22
Those	56	0.08	16	0.01
It	1246	1.82	1775	1.96

Following Table 20, the pronoun *it* is the most frequent pronoun in both corpora with close densities, but the pronoun *this* is used twice more by natives. Consider the examples below.

Ex 5: In other words, language learning is not a singular activity; rather **it** is invested with meaning across different material and discursive contexts. (Native corpus, Norris, 2016)

Ex 6: There seemed to be an obvious and intentional avoidance of taking a power position in news related to the crush, except where it came to responding to Iran's criticism of its mismanagement of the Hajj rituals. In fact, **this** was the one and only example of Saudi Arabia's strong discourse which had been made on Sunday, 27 September 2015. (Non-native corpus, Mohseni, 2015)

Second person pronouns: Writers tend to employ *second person pronouns* to point to readers as specific or broad referents. These pronouns are, as well, more abundantly extant in spoken interactions, with *you* and *your* being by far more widespread in conversation than academic prose (Biber et al., 1999). This is why they introduce a sense of informality to academic articles to forge a relationship with readers. *Second person pronouns* are extremely overused by non-natives than natives

and also vis-à-vis Hyland and Jiang's (2017) corpus. Blagojevic studied some languages which produce writer-based texts and some others select reader-oriented ones. A reason for this difference might be a social-rhetorical framework (Blagojevic, 2004). Therefore, it can be noticed that non-native writers are more reader-oriented. In other words, they tend to engage their reader more than native writers in their articles and build up closer rapport with them.

Table 21. Most frequent Second Person Pronouns in the Non-native and Native Corpora

	Non-native corpus		Native corpus	
	Frequency	Density	Frequency	Density
You	181	0.26	115	0.12
Your	162	0.23	26	0.02

Table 21 represents the details of these pronouns in our corpora. These pronouns are generally deployed by writers to address readers as distinct or general referents, as in the following excerpts.

Ex 7: Self-confidence change: How do think learning English has changed the perception of **your** confidence and **your** learning ability? (Non-native corpus, Mazdayasna, & Noori, 2014)

Ex 8: The GP further drew attention to the fact that 'there are diagnoses that **you** cannot prove, because there is no test that will give **you** a scientific proof'. (Native corpus, Harding, 2014)

Direct questions: One way via which writers explicitly bring readers into their texts is the use of questions which has been given scant attention in previous studies. Raising questions serves as a predominant characteristic of academic discourse, which encourages utilizing queries or problems to decipher in the form of overt interrogatives (Hyland, 2002b). As many contrastive studies of spoken and written language have demonstrated, the use of direct questions is frequent in spoken language because they can invoke personal involvement among participants (e.g. Biber, 1988; Chafe, 1982; Nash, 1986; Tannen, 1982). Chang and Swales (1999) pointed out that *direct questions* may have a strong impact on writing, while *indirect questions* can be conservative. In academic writing, questions invite the reader to follow the writer's stance, seek engagement and lead readers to the writer's viewpoint (Hyland, 2002b). It was revealed that *direct questions* are applied more in non-native articles, depicting non-native writers' tendency towards asking *direct questions* in order not to sound tedious in their articles, and also representing that they are more comfortable than native writers to apply *direct questions* in their writings. As well, they are highly denser than Hyland and Jiang's (2017) corpus.

Ex 9: There is a need to compare and contrast existing secondary accounts, aiming both to correct inaccuracies in and considerably 'fill out' these accounts and assessments, but historical research should also, almost as a sine qua non, make reference to primary sources. **What kinds of primary source evidence are to be consulted, then, and how are primary sources to be sought out, evaluated, and utilized?** This is where the above considerations of the specific scope of ALH come firmly into play. (Native corpus, Smith, 2015)

Ex 10: **What types of CF are used by Iranian EFL teachers teaching in language institutes?** (Non-native corpus, Atai & Asadnia, 2016)

5. Discussion and conclusions

Native and non-native articles were compared in our study in view of the fact that non-native writers who write articles in English consider English as the lingua franca in the academic publishing world, and consequently they should be informed about the linguistic rules in this shared language. By showing these similarities and differences, non-native writers will be able to pursue this language of global academic publication more efficiently.

The overall findings of our study are comparable with previous findings from studies such as Cobb (2003), McCrostie (2008), Petch-Tyson (1998), and Leedham (2015) which indicate a greater use of informality by non-native writers. Shaw (1991) argues that non-native writers seem to be commonly affected by the discipline, genre, and the knowledge about the audience which give rise to their writings problems. Iranian non-natives' failure, therefore, may also be attributed to the nature of the academic article writing genre or the applied linguistics discipline which, as Hyland and Jiang (2017) revealed, has been less informal compared to the three other disciplines in their study. In their perusal of the norms which direct this genre and discipline, they make a lower use of informality, which can also be explicated in light of the particular audience for whom they write their articles. Their articles are usually read internally by a native Iranian audience with their own expectations. Being cognizant of the demands of a stern audience and stringent journal review policies which are not charitable to informality in this genre, Iranian research writers find themselves obliged to employ a smaller degree of informality in their texts.

An additional potential reason for the overall underuse of informality by Iranian non-native speakers can be related to the fact that some of these features, such as *anaphoric pronouns* and *initial conjunctions*, form lexical bundles and follow idiomaticity. Adel and Erman (2012) contend that these combinations tend to emerge more in native than non-native texts because natives possess a larger inventory of types and can make use of various forms. The same holds true for the Iranian authors who write articles in English, despite their highly advanced English proficiency. This is parallel with the conventions in phraseology (Erman, 2009) and findings in the area of lexical bundles (Chen & Baker, 2010). This underuse might also be the result of non-natives' conservatism, as put by Durrant and Schmitt (2009), which compels them to over-depend on more formal and common forms of language.

In agreement with lower use of informality by non-native writers in our result, one possible explanation for differences is that non-native writers tend to focus more on academic rules, textbooks, and formal writing, and they are more concerned with the structure of the English language. A further explanation behind this inappropriate usage is likely attributable to the more-or-less traditional type of instruction supplied in research writing and EAP classes in the Iranian academic context, which places an overemphasis on formal aspects of academic writing, overriding the role that informality can play in article development to engage readers and create more amicable writings. Distinctions between formal and informal writing might be introduced but not adequately taught and practiced. This conventional approach is reinvigorated in many academic writing guidebooks (Coffin et al., 2003). It can be said

that these difficulties in writing appear from a lack of proper awareness and training of these rules which govern academic writing in English.

Although our research indicates that we are witnessing intense employment of certain informality elements by native authors, it is substantial to avoid overemphasizing the range of the usage. There are influential promoters of genres including scholars who have financed eons into procuring their occult norms (Hyland & Jiang, 2017); hence, they judiciously resort to informal features. Native authors also apply these features on the ground that they unswervingly have to do with the rhetorical drives of the writing and ongoing alterations to rules of interpersonal coaxing.

One consequence of our results is the possible difficulties these differences create for learners and apprentice writers. Although informality might be regarded as a benefit to skilled essayists, it can render the liaisons the writer is struggling to generate with readers more intricate and swells to the difficulty of developing compositions by novice writers. It is crucial to adopt a more pronounced method which stresses teaching formality/informality, putting more weight on using authentic texts, including articles written by adroit native writers. Armed with this knowledge and subjected to a discovery-oriented writing approach, non-native researchers can autonomously unearth the uses of these features in academic prose and then translate them in their advanced writings to enhance their participation in the academia, boosting their chances of publishing articles in more esteemed English-medium journals.

Corpus-based instruction can expose the areas of overuse, underuse, or misuse (Flowerdew, 2001; Granger & Tribble, 1998; Hyland, 2001; Ragan, 2001). Specialized computerized corpora can go a long way to provide non-native researchers with concordances that display how informality functions in academic writing and bring to their attention the areas which are prone to deviations from native speakers' norms.

Our outcomes can offer ramifications meant for academic writers to become aware of different conventions governing this kind of genre (Adel & Erman, 2012; Chen & Baker, 2010). Informal features are used in dissimilar ways in native and non-native articles, which makes it important for teachers and ES/AP course designers to recognize this for instruction. These findings unravel what discourse features should be taught in the classrooms, and more certainly, what should be embedded into the EAP courses and textbooks. Academic writing materials developers should benefit from authentic research articles to showcase actual instances of informal language use, or they can include instruction on how to employ specific text analysis softwares, such as AntConc, so that learners can find out on their own how to correctly incorporate informality in their writings.

The findings of the study can also make non-native academic writers aware of the conventions of their English articles and help them to be careful in using different informal features to write their articles in a more natural way. A further strength of this study might be that it contributes to the improvement of the ability to understand the language of written academic discourse. Therefore, native and non-native writers can improve their articles through developing conscious knowledge of

the use and functions of informal features. Further, knowledge of the informal features might help them to understand and create an article with greater ease. The results can also inform those students who want to improve their writing process for a more successful academic career. They are suggested to devote a greater attention to the significant parts of the articles which are highlighted by informal features.

This study was a bid to highlight some of the challenges that non-native writers of English wrestle with in using informality in research articles. We think much more research has to be executed to gain more insights into this slippery concept. Especially, there ought to be more contrastive cross-linguistic and cross-disciplinary studies to understand the impacts of transfer and disciplinary variation on this type of advanced writing.

Due to the fact that research on informality in academic writing is still in its infancy, we must take heed of making strong generalizations about the dissimilarities between natives and non-natives as we only analyzed ten features in the framework which is not inclusive of its true essence. At the heart of informality also reside some other components, such as lexis, grammar, and pragmatic aspects, to which we and other researchers so far have been seemingly oblivious. As a result, further research needs to be performed to bridge this gap.

References

- Adel, A. (2008). Metadiscourse across three varieties of English: American, British and advanced learner English. In U. Connor, E. Nagelhout, & W. V. Rozycki (Eds.), *Contrastive rhetoric: Reaching to intercultural rhetoric* (45-62). Amsterdam: Benjamins.
- Adel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for Specific Purposes*, 31, 81-92.
- Alipour, M. & Matouri, H. (2017). Comparative study of reflexive metadiscourse in applied linguistics research articles published in international journals and in Iranian national journals. *ARTESOLESP E-Journal*, 7(1), 15-28.
- Anthony, L. (2011). *AntConc* 3.4.3. <http://www.laurenceanthony.net/software.html>
- Atai, M. R., & Asadnia, F. (2016). The prestigious world university on its homepage: The promotional academic genre of overview. *Iranian Journal of Applied Linguistics*, 19(1), 1-34.
- Atkinson, D. (1999). *Scientific discourse in sociohistorical context: The philosophical transactions of the Royal Society of London, 1675–1975*. Mahwah, NJ: Lawrence Erlbaum.
- Attarn, A. (2014). Study of metadiscourse in ESP articles: A comparison of English articles written by Iranian and English native speakers. *International Journal of Learning, Teaching and Educational Research*, 5(1), 63-71.
- Aufa, F. (2014). Should native speaker norms be taken into account?: A perspective in teaching EIL. *SELT*, 11, 152-258.
- Belcher, D. (2007). Seeking acceptance in an English-only research world. *Journal of Second Language Writing*, 16, 1–22.
- Bennet, K. (2009). English academic style manuals: A survey. *Journal of English for Academic Purposes*, 8, 43-54.
- Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.

- Biber, D. (1995). *Dimensions of register variation: A cross-linguistic comparison*. Cambridge: Cambridge University Press.
- Biber, D., Conrad, S., & Reppen, R. (1999). *Corpus linguistics. Investigating language structure and use*. Cambridge: Cambridge University Press.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow: Pearson.
- Blagojevic, S. (2004). Metadiscourse in academic prose: A contrastive study of academic articles written in English by English and Norwegian native speakers. *Studies about Languages*, 5, 60-67.
- Brown, P., & Levinson, S. (1978). *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.
- Celik, S. (2006). A concise examination of the artificial battle between native and non-native speaker teachers of English in Turkey. *Kastamonu Education Journal*, 14(2), 371-376.
- Chafe, W. (1982). 'Integration and involvement in speaking, writing, and oral literature'. In T. Deborah (Ed.), *Spoken and written language: Exploring orality and literacy* (pp. 35-54). Norwood, NJ: Ablex.
- Chafe, W. (1986a). Evidentiality in English conversation and academic writing. *Evidentiality: The linguistic coding of epistemology*, 20, 261-272.
- Chafe, W. (1986b). Writing in the perspective of speaking. In C. Cooper, & S. Greenbaum (Eds.), *Studying writing: Linguistic approaches*. London: Sage Publications.
- Chang, Y.-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). London: Longman.
- Chen, Y. H., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning and Technology*, 14(2), 30–49.
- Cobb, T. (2003). Analyzing later interlanguage with learner corpora: Quebec replications of three European studies. *Canadian Modern Language Review*, 59(3), 393-423.
- Coffin, C., Curry, M., Goodman, S., Hewings, A., Lillis, T., & Swann, J. (2003). *Teaching academic writing: A toolkit for higher education*. London: Routledge.
- Duff, P. A. (2015). Transnationalism, multilingualism, and identity. *Annual Review of Applied Linguistics*, 35, 57-80.
- Durrant, P. (2013). Discipline and level specificity in university students' written vocabulary. *Applied linguistics*, 35(3), 328-356.
- Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *IRAL*, 47, 157-177.
- Erman, B. (2009). Formulaic language from a learner perspective: What the learner needs to know. In B. Corrigan, H. Quali, E. Moravcsik, & K. Wheatley (Eds.), *Formulaic language* (pp. 27–50). Amsterdam: John Benjamins.
- Evenson, L. S., & Rygh, I. L. (1988) Connecting L1 and FL in discourse-level performance analysis. *Papers and Studies in Contrastive Linguistics*, 22, 133-178.
- Faghih, E., & Rahimpour, S. (2009). Contrastive rhetoric of English and Persian written texts: Metadiscourse in applied linguistics research articles. *Rice Working Papers in Linguistics*, 1, 92-107.
- Fairclough, N. (2001). *Language and power*. Essex: Longman.
- Field, Y., & Yip, L M. O. (1992). A comparison of internal cohesive conjunction in the English essay writing of Cantonese speakers and native speakers of English. *RELC Journal*, 23(1), 15-28.
- Florence, L. P. (2012). Advantages and disadvantages of native- and non-native Englishspeaking teachers: Students perceptions in Hong Kong. *TESOL Quarterly*, 46(2), 280- 305.

- Foster, J. (2005). *Effective writing skills for public relations*. London: Kogan Page.
- Flowerdew, L. (2001). The exploitation of small learner corpora in EAP materials design. In M. Ghadessy, A. Henry, & R. L. Roseberry (Eds.), *Small corpus studies and ELT* (pp. 363–379). Amsterdam: John Benjamins.
- Ghafournia, N., & Afghari, A. (2013). Exploring the relationship between learning strategies, academic disciplines, and reading comprehension test performance. *Iranian Journal of Applied Linguistics*, 16(2), 21–51.
- Granger, S., & Tribble, C. (1998). Learner corpus data in the foreign language classroom: Form-focused instruction and data-driven learning. In S. Granger (Ed.), *Learner English on computer* (pp. 199–211). London: Longman.
- Geisler, C., Kaufer, D. S., & Steinberg, E. R. (1985). The unattended anaphoric ‘This’: When should writers use it? *Written Communication*, 2, 129–155.
- Gilbert, G. N., & Mulkay, M. (1984). *Opening Pandora's box: A sociological analysis of scientists' discourse*. Cambridge: Cambridge University Press.
- Halliday, M. A. K. (1985). *Spoken and written language*. Oxford: Oxford University Press.
- Hardie, A. (2014). Log Ratio: An informal introduction. Retrieved from <http://cass.lancs.ac.uk/?p=1133>
- Harding, L. (2014). Towards a theory of diagnosis in second and foreign language assessment: Insights from professional practice across diverse fields. *Applied Linguistics*, 36(2), 236–260.
- Heylighen, F., & Dewaele, J. M. (1999). *Formality of language: Definition, measurement and behavioral determinants*. Internal Report. Center “Leo Apostel”: Free University of Brussels.
- Howarth, P. (1998). The phraseology of learners’ academic writing. In A. Cowie (Ed.), *Phraseology* (pp. 161–186). Oxford: Clarendon Press.
- Hyland, K. (2001). ‘Bringing in the reader: Addressee features in academic articles,’ *Written Communication*, 18(4), 549–74.
- Hyland, K. (2002a). Authority and invisibility: Authorial identity in academic writing. *Journal of Pragmatics*, 34, 1091–1112.
- Hyland, K. (2002b). What do they mean? Questions in academic writing. *Text-The Hague Then Amsterdam Then Berlin*, 22(4), 529–558.
- Hyland, K. (2004). *Disciplinary discourses: Social interactions in academic writing*. Ann Arbor, MI: University of Michigan Press.
- Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, 7(2), 173–191.
- Hyland, K. (2009). English for professional academic purposes: Writing for scholarly publication. In D. A. Belcher, & A. Arbor (Eds.), *English for specific purposes in theory and practice* (pp. 83–105). MI, USA: University of Michigan Press.
- Hyland, K. (2009). Constraint versus Creativity: identity in academic writing. In M. Gotti (Ed.) *Commonality and Individuality in Academic Discourse* (pp. 25–520). Frankfurt: Peter Lang.
- Hyland, K., & Anan, E. (2006). Teacher’s perceptions of error: The effects of first language and experience. *System*, 34, 509–519.
- Hyland, K., & Jiang, F.K. (2017). Is academic writing becoming more informal?. *English for Specific Purposes*, 45, 40–51.
- Hofstede, G. (1991). *Cultures and organizations: software of the mind*. London: McGraw-Hill.
- Hundt, M., & Mair, C. (1999). “Agile” and “Uptight” genres: The corpus-based approach to language change in progress. *International Journal of Corpus Linguistics*, 4(2), 221–242.
- Ivanic, R., & Camps, D. (2001). I am how I sound. Voice as self-representation in L2 writing. *Journal of Second Language Writing*, 10, 3–33.
- Jenkins, J. (2000). *The phonology of English as an international language*. Oxford: Oxford University Press.

- Jenkins, J. (2000). The spread of EIL: A testing time for testers. *ELT Journal*, 60(1), 42-50.
- Jenkins, J. (2003). *World Englishes: A resource book for students*. Oxford: Oxford University Press.
- Jenkins, J. (2011). Accommodating (to) ELF in the international university. *Journal of Pragmatics*, 43, 926–936.
- Jiang, K., & Hyland, K. (2015). “The fact that”: Stance nouns in disciplinary writing. *Discourse Studies*, 17(5), 529-550.
- Johnston, J. E., Berry, K. J., & Mielke, P. W. (2006). Measures of effect size for chi-squared and likelihood-ratio goodness-of-fit tests. *Perceptual and Motor Skills*, 103, 412-414.
- Kahkesh, M., & Alipour, M. (2017). A comparative study of metadiscourse markers in English and Persian university lectures. *Research in Applied Linguistics*, 8, 125-135.
- Kaplan, R. B. (1966). Cultural thought patterns in inter-cultural education. *Language learning*, 16(1-2), 1-20.
- Keshavarz, M. H., Kheirich, Z. (2011). Metadiscourse elements in English research articles written by native English and nonnative Iranian writers in Applied Linguistics and Civil Engineering. *Journal of English Studies*. 1(3), 3-15.
- Kuo, C. (1999). The use of personal pronouns: Role relationships in scientific journal articles. *English for Specific Purposes*, 18(2), 121–138.
- Leedham, M. (2015). *Chinese students' writing in English: Implications from a corpus-driven study*. Oxford: Routledge.
- Lillis, T., & Curry, M. (2010). *Academic writing in a global context: The politics and practices of publishing in English*. London: Routledge.
- Mair, C. (1998). Corpora and the study of the major varieties of English: Issues and results. *The major varieties of English: Papers from Mavén*, 97, 139-157.
- Martinez, I. A. (2000). Discrepancias entre las concepciones de hispanoparlantes del artículo de investigación en inglés y las características del mismo según el análisis del género. Actas II Congreso Iberoamericano de Educación en Ciencias Experimentales. Córdoba, Argentina: CD-Rom.
- Martinez, I. A. (2005). Native and non-native writers' use of first person pronouns in different sections of biology research articles in English. *Journal of Second Language writing*, 14, 174-190.
- Mauranen, A. (2009). Introduction. In A. Mauranen, & E. Ranta (Eds.), *English as a lingua franca: Studies and findings* (pp. 1–9). Newcastle: Cambridge Scholars Publishing.
- Mazdayasna, G., & Noori, M. (2014). Developing a profile of learning needs of Iranian undergraduate students of English Language and Literature. *Iranian Journal of Applied Language Studies*, 6(2), 27-58.
- McCrostie, J. (2008). Writer visibility in EFL learner academic writing: A corpus-based study. *Icame Journal*, 32, 97-114.
- Milton, J., & Tsang, E. S. C. (1993). A corpus-based study of logical connectors in EFL students' writing: directions for future research. In R. Pemberton, & E.S.C. Tsang (Eds.), *Studies in lexis* (pp. 2 15-246). Hong Kong: The Hong Kong University of Science and Technology.
- Mirzapour, F., RasekhMahand, M. (2012). Hedges and boosters in native and non-native library and information and computer science research articles. *The Southeast Asian Journal of English Language Studies*, 18(2), 119-128.
- Mohseni, A. (2015). Investigating the relationship between teacher's thinking vs. feeling personality type and Iranian pre-intermediate EFL learners' speaking skill. *Iranian Journal of Applied Language Studies*, 5(2), 123-136.
- Moreno, A. (2010). Researching into English for research publication purposes from an applied intercultural perspective. In M. Garrido, J. Palmer-Silveria, & I. Fortanet-Gómez (Eds.), *English for professional and academic purposes* (pp. 57–71). New York: Rodopi.

- Nash, W. (1986). *English usage: A guide to first principles*. London: Routledge & Kegan Paul.
- Norris, J. M. (2016). Current uses for task-based language assessment. *Annual Review of Applied Linguistics*, 36, 230-244.
- Nwogu, K. N. (1997). The medical research paper: Structure and functions. *English for Specific Purposes*, 16(2), 119–138.
- Ozturk, U. & Atay, D. (2010). Challenges of being a non-native English teacher. *Educational Research*, 1(5), 135-139.
- Petch-Tyson, S. (1998). Writer/reader visibility in EFL written discourse. In S. Granger (Ed.), *Learner English on computer* (pp. 107-118). London: Addison Wesley Longman.
- Ragan, P. (2001). Classroom use of systemic functional small learner corpus. In M. Ghadessy, A. Henry, & R. L. Roseberry (Eds.), *Small corpus studies and ELT* (pp. 207–236). Amsterdam: John Benjamins.
- Rayson, P. (2008). From key words to key semantic domains. *International Journal of Corpus Linguistics*, 13(4), 519-549.
- Scott, M. (2001). Comparing corpora and identifying key words, collocations, frequency distribution through the WordSmith Tools suite of computer programs. In M. Ghadessy, A. Henry, & R. L. Roseberry (Eds.), *Small Corpus Studies and ELT* (pp. 47–67). Amsterdam: John Benjamins.
- Seone, E. (2013). On the conventionalisation and loss of pragmatic function of the passive in Late Modern English scientific discourse. *Journal of Historical Pragmatics*, 14(1), 70-99.
- Seone, E., & Loureiro-Porto, L. (2005). On the colloquialization of scientific British and American English. *ESPAcross Cultures*, 2, 106-118.
- Shaw, P. (1991). Science research students' composing processes. *English for Specific Purposes*, 10, 189–206.
- Sifakis, N. C. (2014). Teaching pronunciation in the post-EFL era: Lessons from ELF and implications for teaching education. In J. D. M. Agudo (Ed.), *English as a foreign language teacher education: Current perspective and challenges* (pp. 127-146). Amsterdam: Rodopi.
- Smith, R. (2015). Building 'applied linguistic historiography': Rationale, scope, and methods. *Applied Linguistics*, 37(1), 71-87.
- Straus, P. (2017). It's not the way we use English: Can we resist the native speaker stranglehold on academic publications? *Publications*, 5, 1-7.
- Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Swales, J. M., & Feak, C. B. (2012). *Academic writing for graduate students: A course for nonnative speakers of English*. Ann Arbor, MI: University of Michigan Press.
- Tang, R., & John, S. (1999). The "I" in identity: Exploring writer identity in student academic writing through the first person pronoun. *English for Specific Purposes*, 18, 23–39.
- Tannen, D. (1982). Oral and literate strategies in spoken and written narratives. *Language*, 58, 1-21.
- Vakili Latif, S. (2016). Discourse and ideology variation: A critical functional approach to Mina stampede news reports. *Iranian Journal of Applied Linguistics*, 19(1), 207-231.
- Yeganeh, M. T., Heravi, I. M., & Sawari, A. (2015). Hedge and booster in newspaper articles on Iran's presidential election: A comparative study of English and Persian articles. *Procedia-Social and Behavioral Sciences*, 192, 679-683.
- Zarei, G. R., & Mansoori, S. (2011). A Contrastive study on metadiscourse elements used in humanities vs. non humanities across Persian and English. *English Language Teaching*, 4(1), 42-50.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the Journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (CC BY-NC-ND) (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).