Hedges in English for Academic Purposes: A Corpus-based study of Iranian EFL learners

Hossein Khazaee, Ph.D. Candidate, Department of English, Science and Research Branch, Islamic Azad University, Tehran, Iran
khazaee.h.1980@gmail.com
Parviz Maftoon, Associate Professor, Department of English, Science and Research Branch, Islamic Azad University, Tehran, Iran
pmaftoon@srbiau.ac.ir
Parviz Birjandi, Professor, Department of English, Science and Research Branch, Islamic Azad University, Tehran, Iran
pbirjandi@srbiau.ac.ir
Ghafour Rezaie Golandouz, Assistant Professor, Department of English, Garmsar Branch, Islamic Azad University, Garmsar, Iran
gh.rezaie@iau-garmsar.ac.ir

Abstract
Hedges, as tools to express tentativeness and doubt, have been studied in plenty of research papers in the Iranian EFL research setting. However, their use in a learner corpus, portraying Iranian learner English, is in need of more research attention. With this end in view, this study aimed at investigating how Iranian EFL learners who have majored in English-related fields in Iran deployed hedges in their academic, expository essays. This study was conducted through running the corpus analysis software MonoConc Pro-Semester version 2.2 on the electronically compiled Iranian Corpus of Learner English, totaling 436,035 words. Automatic and manual analyses suggested that hedges comprised only 7.4% of the total metadiscourse in the Iranian Corpus of Learner English, with 0.68 occurrences per 1,000,000 words. In a comparable native corpus, a sub-corpus of the British Academic Written English, hedges were used with 1.43 occurrences per 1,000,000 words (21% of the total metadiscourse in the corpus). Log-likelihood statistical analysis confirmed statistically significant differences between the two corpora in terms of the use of hedges, with underuse of hedges in the Iranian academic, expository essays relative to the English natives' essays. Implementations of the results for English academic writing instruction including genre-based, explicit teaching of hedges through data-driven techniques with the aid of tools such as AntConc software and corpora such as the BAWE are considered.

Keywords: Academic writing, Corpus Builder Software, hedges, Iranian corpus of learner English, learner corpora

Introduction

English for academic purposes (EAP), as a strand of English for specific purposes (ESP), aims at equipping English as a second or foreign language (ESL/EFL) learners with the academic language knowledge, critical thinking, and cultural skills that are necessary for attaining full academic potentials. Among the issues that have been the focus of EAP researchers during the last decades was the compilation of electronic corpora as sources of information concerning EFL learners' use of hedges in genres such as academic essays, academic research papers, Ph.D. dissertations, and so forth.

Hedges are linguistic tools that are used to express tentativeness, doubt, and possibility (Hyland, 2017). Hedges have received considerable attention in studies on both spoken discourse

and written academic discourse. In spoken discourse, for example, their role "in qualifying categorical commitment and facilitating discussion" (Hyland, 1996, p. 433) has been investigated by many researchers in fields such as conversation analysis (e.g., Varma & Tan, 2015). In Written academic discourse, plenty of studies have been conducted on the significant role that hedges play in writers' need to propose their unverified suggestions with care (e.g., Dontcheva-Navratilova, 2016; Plappert, 2019, among others). Hedges have also been studied as tools for indexing fuzziness in discourse (e.g., Channell, 1994), as metadiscourse markers (e.g., Hyland, 2005; Vande Kopple, 1985), and as a means of indexing detachment between a speaker and what he/she is saying (e.g., Ahn & Yap, 2015).

A glance through the literature of EAP studies reveals that the use of hedges in written academic discourse has been focused on in various theoretical models (e.g., Burrough-Boenisch, 2005; Hyland 1996, 1998, 2005; Markkanen & Schröder, 1997; Mauranen, 1997). Hyland (1998, pp. 39-50) categorizes the linguistic approaches that deal with hedges into three classes, namely (a) "speech acts and interpretive maxims," (b) "epistemic modality," and (c) "metadiscourse." In Hyland's (1998) conceptualization of hedging, "hedges are the means by which writers can present a proposition as an opinion rather than a fact: items are only hedges in their epistemic sense, and only when they mark uncertainty" (p. 5). Based on Hyland (1998), hedges are categorized into lexical and strategic categories. Lexical hedges are again grouped into modal auxiliaries, epistemic lexical verbs (including epistemic judgment and evidential verbs), epistemic adverbs, epistemic adjectives, and epistemic nouns. Strategic hedges are grouped into reference to limited knowledge, reference to limitation of model, theory, or method, and reference to experimental limitations.

Analyses of language corpora have revealed some of the characteristics of hedges in textbooks (Bouhlal, Horst, & Martini, 2018), economic forecasting (Resche, 2015), scientific articles (Vass, 2017), abstracts (Li & Pramoolsook, 2015), medical discourse (Martikainen, 2018), and so forth. A corpus is an electronic collection of naturally occurring texts that is used in language-related studies (Hunston, 2002). Hedges have also been studied in plenty of corpus-informed, corpus-based, or corpus-driven learner language investigations (e.g., Larsson 2017; Sun & Hu, 2020, among others). Learner language is defined as "the oral or written language produced by learners" (Ellis & Barkhuizen, 2005, p. 4). In corpus-informed investigations (e.g., Ackermann, Biber, & Gray, 2011; Hawkey & Barker, 2004), researchers analyze a learner corpus for information concerning the existence, non-existence, or probable errors with reference to a specific linguistic feature (Callies, 2015). As Tognini-Bonelli (2001) mentions, in corpus-based analysis procedures (e.g., Hawkins & Filipović, 2012), the researcher initiates the analysis of a learner corpus with a list of pre-categorized linguistic features, while in corpus-driven analyses (e.g., Wulff & Gries, 2011), results emerge out of the learner corpus without any linguistic pre-specifications.

Taking the importance of hedging as a resource for expressing uncertainty in academic discourse, lack of analysis of hedges in a learner corpus, portraying academic, expository essays of Iranian EFL learners, is really felt. Contributing to this line of investigation, this study, with the aid of the electronically compiled Iranian Corpus of Learner English and a sub-corpus of the British Academic Written English (BAWE), aims at finding the probable overuses or underuses of hedges in Iranian EFL learners' academic, expository essays from a metadiscursive aspect.

As a learner corpus research (LCR) investigation, based on Hyland's (2005) interpersonal metadiscourse model, hedging is considered as a sub-category of metadiscourse in the present paper. It should be mentioned that there is not a consensus among researchers in including hedges as metadiscourse (e.g., Ädel, 2006, 2010; Mauranen, 1993a, 1993b, 2010). Hyland's (2005)

metadiscourse model includes 101 hedging items that consist of categories such as lexical and strategic hedges including epistemic modal auxiliaries (e.g., *could*, *couldn't*, *may*, etc.), epistemic lexical verbs (e.g., *appear*, *argue*, *assume*, *claim*, etc.), epistemic adverbs (e.g., *about*, *almost*, *apparently*, *approximately*, etc.), and so forth. With this end in view, the research questions of this study are:

Q1. From a metadiscursive aspect, what are the most frequently used hedges in the academic, expository essays of Iranian intermediate EFL learners who have majored in English-related fields in Iran?

Q2. From a metadiscursive aspect, what what are the most frequently used hedges in the English-native, linguistics students' academic, expository essays?

Q3. Concerning the frequencies of the use of hedges, are there any statistically significant differences (overuses and underuses) between the Iranian intermediate EFL learners' and English-natives' academic, expository essays?

Q4. If any differences (overuses and underuses) exist between the Iranian intermediate EFL learners' and English-natives' academic, expository essays in terms of using hedges, what are the possible reasons of such overuses or underuses?

Q5. If any differences (overuses and underuses) exist between the Iranian intermediate EFL learners' and English-natives' academic, expository essays in terms of using hedges, are there any mechanisms to alleviate such differences?

Literature Review

Previous Corpus Studies on the Use of Hedges

A glance through the literature of corpus studies reveals plenty of large-sized corpus research papers on the important role that hedges play in the academic writings of both native and non-native speakers of English. Such studies refer to the differences in the use of hedges between English-native writers and writers from other first language (L1) backgrounds. Most of the studies in the Iranian EFL context, however, focus on the use of hedges in small-sized corpora of genres such as academic research papers, Ph.D. dissertations, and examples of LCR studies on academic essays, especially expository essays that Iranian university-level EFL learners write, are rarely found. With this end in view, following are a number of recent corpus studies on the use of hedges in the Iranian EFL context.

Tahririan and Shahzamani (2009) compared the frequency of the use of hedges between English and Persian social, economic, and political newspaper editorials. The results indicated that English newspaper editorials employed more hedges than Persian ones. Moreover, English political editorials were slightly more hedged than the economic and social ones. With regard to the Iranian context, economic editorials were slightly more hedged than the political and social ones.

Jalilifar, Shooshtari, and Mutaqid (2011) investigated the impacts of the explicit instruction of hedges on the reading comprehension of ESP materials of Iranian English-related university students. The results indicated the facilitative effects of explicit instruction in recognizing hedging devices that improved the language proficiency of the students and therefore improved their reading comprehension scores.

Falahati (as cited in Rasti, 2011) investigated the use of hedges in 24 research articles, written in English by English-native writers and Persian EFL learners. His quantitative analysis revealed that English writers employed hedges more frequently than their Iranian counterparts.

In a study with rather contradictory findings compared with the literature of research on the use of hedges in the Iranian EFL context, Abdi and Behnam (2014) investigated the use of

hedges in the abstracts of medical articles written by English and Persian writers. They found that there were not any statistically significant differences between Iranian and American writers in terms of the use of hedges.

Rezanejad, Lari, and Mosalli (2015) compared the frequency of the use of hedges in different sections of research papers written by Iranian authors and English-native authors. The results showed an underuse of hedges by Iranian authors, and there was a statistically significant difference between these two groups of authors' use of hedges.

Azarbad and Ghahraman (2018) investigated the distribution of functions and forms of hedges in the English and Persian abstracts of master's theses, written by Iranian students. The results showed that there was a statistically significant difference between English and Persian texts concerning the use of hedges, with underuse of hedges in the Persian ones. Azarbad and Ghahraman referred to the degree of awareness, purpose, cultural background of the learners, and so forth as determining factors with regard to the observed differences of the use of hedges in their investigation.

Sabzevar, Haghverdi, and Biriya (2020) explored the use of epistemic adverbs in academic essays written by English-native speakers and Iranian EFL learners. The findings revealed the underuse of hedges among Iranian EFL writers.

Weisi and Asakereh (2020) investigated the impact of gender and nativeness on the use of hedges in the discussion section of applied linguistics research papers, written by natives of English and Iranian EFL research writers. The results indicated that there was a statistically significant difference between Iranian male and female research writers in terms of the frequency of use of hedges, that is, gender had a determining impact on the use of hedges. Moreover, the results showed that the discussion sections of applied linguistics research papers, written by English-native research writers, were more hedged than those written by their Iranian counterparts.

Method

Corpus

Use of corpus analysis software and online corpus-related websites (with in-built engines to extract and analyze features such as concordancing lines, hypertext, N-grams, PoS tags, range, etc.) in language teaching (for purposes such as data-driven learning (DDL), has become a common practice among LCR investigators. Such software and websites are also used in EAP instruction courses and research. Among the software and websites for DDL practices and analyses of corpora, Anthony's (2017) AntConc series, Tsukamoto's (2002) KWIC, Barlow's (2017) MonoConc Pro, and Cobb's (2016) the Compleat Lexical Tutor can be named here (a complete and up-to-date list of corpus analysis software and online corpus-related websites can be found at https://corpus-analysis.com).

The common technicality issue among most of these software and websites is their user-friendly interfaces that users can search different aspects of a language within them and get their intended results. Even teachers can apply them in their DDL practices and obtain excellent results in terms of helping learners to reach learner autonomy and flooding them with input from real language use. The search interface of AntConc software version 3.5.7, from Anthony's (2019) AntConc series, is presented in figure 1.

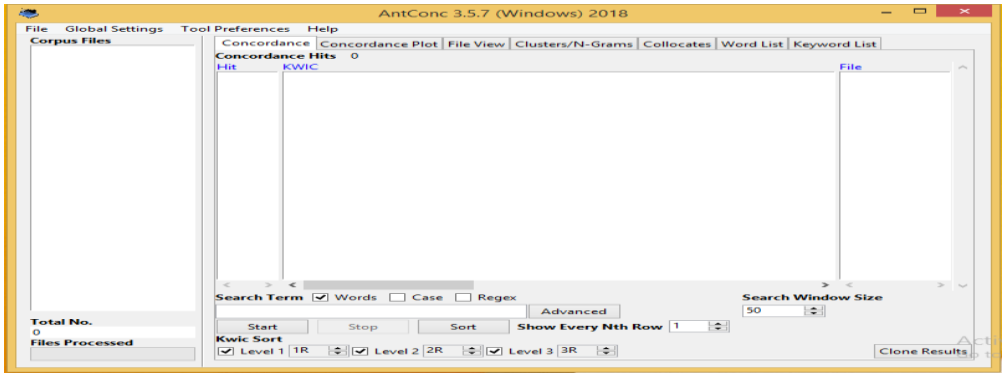

Figure 1. *Search Interface of AntConc*

As is seen in figure 1, the researcher or the teacher can write a language term in the search space and obtain different results including the frequency of the searched term in the intended corpus, keywords, common collocations that are used with the intended term, and so forth. Besides these software and websites, there are numerous online and archived corpora around the world now. Examples are the British National Corpus, the Corpus of Contemporary American English, and so forth. Learner corpora are no exceptions in this regard, and now there are plenty of learner corpora around the world, which have been designed by researchers from various L1 backgrounds (a complete and up-to-date list of learner corpora around the world can be found at https://uclouvain.be/en/research-institutes/ilc/cecl/learner-corpora-around-the-world.html).

A very important issue in compiling learner corpora is the design criteria that are followed in them (Atkins, Clear, & Ostler, 1992). Without such criteria, as Granger (2012) points out, we will face a haphazard compilation of learner data that cannot be called a leaner corpus. Ellis (2008) names learner and task variables as important criteria in designing leaner corpora. Granger (2003) follows a rigorous account of more than 20 learner and task variables in designing the International Corpus of Learner English, which has been frequently referred to as a well-established paradigm for corpus designers to follow along (Ellis & Barkhuizen, 2005; Lozano & Mendikoetxea, 2013). An illustration of the design criteria of the International Corpus of Learner English is provided in figure 2 Below:

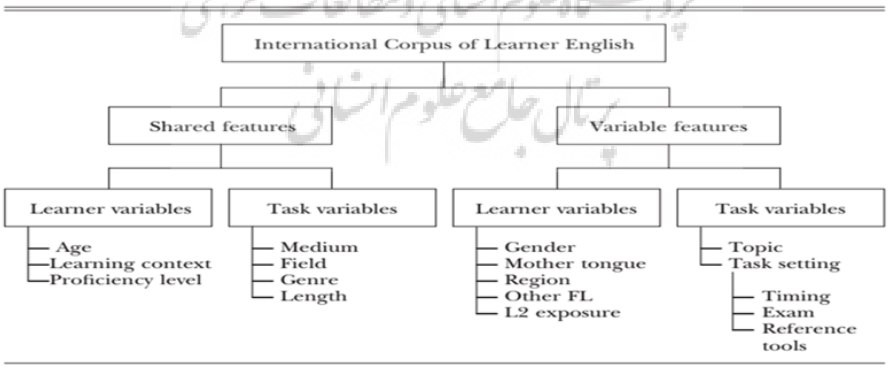

Figure 2. *Task and Learner Variables in the International Corpus of Learner English*

Compilation and Features of the Iranian Corpus of Learner English

The Iranian Corpus of Learner English, designed for the sake of the Ph.D. dissertation of the first author of the current paper, was compiled between March 2017 and January 2018. It

includes 436,035 words and consists of 1,744 expository essays that were written by male and female Persian native speakers, who attended English-related fields at both State Universities and Islamic Azad University branches of Iran. These EFL learners were in junior and senior years of their education, and before the compilation of the corpus, they had passed courses such as paragraph development and essay writing in the Iranian curriculum on English-related fields. With regard to the variable of age, these learners were between 21 and 35 (Mean= 28). All of the participants were students of the same university for at least four consecutive semesters.

The 50-min EF Standard English Test (EFSET) was used to check the proficiency level of these learners, and before the compilation of the corpus, learners' essays whose proficiency levels were intermediate (EFSET's B1) were put aside for further analyses. These task and learner criteria were checked through a learner profile form that the learners filled in and together with their essays sent to the first author of the current paper. Part of the corpus data was obtained from two female colleagues whose research also included expository essays of Iranian intermediate university-level learners who were majoring in English related fields in Iran.

The learners were asked to write 250-word expository essays on the topics of the book *Expository Eureka: Model Expository Essays for Today's Secondary School Students* (Tham, 2013). They were also asked to submit copies of their hand-written or typed essays to the first author of this study. The essays were gathered as untimed, authentic texts. This means that the learners had access to reference tools such as dictionaries, and no corrections were made to the errors in the essays.

After gathering the essays, the handwritten ones were scanned and converted to the JPEG format. Next, through applying the optical character recognition software (OCR), they were converted into machine-encoded texts (i.e., Microsoft word document files). Part of the data that was given by the two female colleagues included typed essays and did not require this step. In the next step of the compilation of the corpus and through applying the AntFileConverter freeware (Anthony, 2017), the essays were converted into PDF and then plain texts. Finally, the files were uploaded in the corpus builder engine at http://www.lextutor.ca/ (Cobb, 2016), and the Iranian Corpus of Learner English was built and ready for the analysis of use of hedges.

Compilation and Features of the BAWE Sub-corpus

The BAWE corpus, totaling 6,506,995 words, includes texts in genres such as case studies, empathy writings, essays, and so forth in four academic domains, namely arts and humanities, social sciences, life sciences, and physical sciences. The texts in the corpus have been compiled from university students across four levels of study in 35 disciplines. For the sake of comparison with the Iranian Corpus of Learner English, 44 texts from a sub-part of the BAWE corpus including English-native, bachelor of arts (BA), linguistics students' expository essays, totaling 92,984 words, was used. These essays were downloaded as resource number 2,539 from the University of Oxford text archive at http://ota.ahds.ac.uk/headers/2539.xml and compiled as a corpus through the corpus builder engine at http://www.lextutor.ca/ (Cobb, 2016).

Procedure

With the aid of the text analysis and concordance software MonoConc Pro-Semester version 2.2. (Barlow, 2017), the Iranian Corpus of Learner English and the sub-part of the BAWE corpus were electronically searched for 101 hedges (appendix of Hyland, 2005). Later, based on justifications to distinguish metadiscoursal hedges from non-hedging elements proposed by Hyland in his various publications (e.g., 1996, 1998, 2005), all seemingly metadiscoursal hedges were manually checked by the first writer of the present paper to ensure they were

performing the role of hedges. This second step (manual checking) is of utmost importance in any LCR study since language items can play various roles in discourse, and as Hyland (2005, p. 218) asserts, every instance of any metadiscoursal element (including hedges per se) "should therefore be studied in its sentential co-text." To clarify the multi-functionality of hedging items, an example of *seems* is presented here.

1.I like and respect him. But recently I've been thinking a lot about a person I see once a week at a club. She *seems* very masculine, with short hair, jeans and heavy boots and she has a sort of butch look I can't describe. **(British national corpus, CH1, W_newsp_tabloid)**

2.The priority given to each of the above rules has depended on the temperament of the president in power but it *seems* that the president's role has been much more forceful, interventionist and political than a reading of the constitution would imply. **(ICLE-brsur1.cor-codes 16-33)**

As is seen, in (a) seems shows ideational (non-hedging) content, while in (b), it functions as epistemic hedge marker.

Results

In the first step of analysis (automatic and then manual), raw frequencies of hedges were obtained from both corpora. As raw frequencies in corpora with different sizes provide uncomparable results, in the second step, these raw frequencies were normalized per 1,000,000 words. Normalizing per 1,000,000 words can be done manually with the formula (raw frequency x 1,000,000) ÷ number of words in the corpus, or electronically, through a number of websites designed for such purposes. Here and in this study, normalizing (per 1,000,000 words) was conducted with the use of the normalizing calculator at http://www.thegrammarlab.com/?p=160. The raw and normalized frequencies of the analyses are depicted in table 1.

Table 1. *Hedges in the Iranian Corpus of Learner English Versus the BAWE Sub-corpus*

Hedges	Raw frequencies	Normalized frequencies	% of total Metadiscourse
ICLE	2,964	0.68	7.4
BAWE Sub-corpus	1,327	1.43	21.0

Note. ICLE stands for the Iranian Corpus of Learner English.

In order to find out whether the observed differences of the use of hedges between the Iranian Corpus of Learner English and the BAWE sub-corpus were statistically significant, table 2 shows the results of the log-likelihood test (*G*-test). The log-likelihood test is a common statistical procedure in LCR that looks at the probable frequency differences between two corpora and analyzes whether a frequency difference is statistically significant or not. The log-likelihood test can be conducted with the aid of many statistical programs such as R. Here and in this study, it was computed with the use of an online calculator (Rayson, 2019) at Lancaster university's center for computer corpus research on language (http://ucrel.lancs.ac.uk/llwizard.html).

Table 2. *Log-likelihood Test Results to Compare the Frequency of Hedges Between the Iranian Corpus of Learner English and the BAWE Sub-corpus*

Item	O1	%1	O2	%2	LL	Bayes
Hedges	2,964	0.68	1,327	1.43	- 452.16	438.98

Note. Based on Rayson (2019), O1 and O2 are the observed frequencies of hedges in the Iranian Corpus of Learner English and the BAWE sub-corpus; %1 and %2 values show normalized frequencies in O1 and O2 per 1,000,000 words. LL indicates the log-likelihood value (G^2); "+" before LL indicates overuse in O1 relative to O2, "-" before LL indicates underuse in O1 relative to O2. Bayes Factor (BIC) indicates effect size: 0-2: not worth more than a bare mention; 2-6: positive evidence against H0; 6-10: strong evidence against H0; > 10: very strong evidence against H0; For negative scores, the scale is read as "in favor of" instead of "against".

As table 2 shows, a statistically significant difference was detected between the two corpora in terms of the use of hedges (G^2>3.8; p < 0.05; with a very large effect size (BIC > 10), with underuse of hedges in the Iranian Corpus of Learner English, relative to the BAWE sub-corpus.

In addition to the above-mentioned analysis results, it should be mentioned that in the Iranian Corpus of Learner English the epistemic modal verbs had the highest frequency among all hedging items with *should* (340 cases of occurrence) as the most frequent one. In the BAWE sub-corpus, too, the epistemic modal verbs had the highest frequency with *may* (162 cases of occurrence) as the most frequently occurring hedging item. The observed frequency of all hedging items in the two corpora are presented in Appendix A of the current study. Below are a number of examples of the use of hedges from both corpora.

3.Finally part time job can make student improve faster but they have to manage their time because it wastes lots of time for studying and they have to be careful. A universal language should be replacing all languages. (Iranian Corpus of Learner English)

4.One must note however, that due to the employers misunderstanding of the proper use and meanings of his/her first words (i.e. mismatches, overextensions, holophrasing...), some may not consider them as the being true language. Finally, the first 50 words may also determine different backgrounds, reflecting the culture into which they were socialized. (The BAWE sub-corpus)

5. In my opinion, I think children from family where both parents work may have some certain advantages. (Iranian Corpus of Learner English)

6. One of the other things that it might make you sick of working when you are a student is finding some jobs that.... (Iranian Corpus of Learner English)

7. It is likely that people who are more powerful will not receive compliments, therefore complimenting maybe a form of subordinating women into a less powerful position in society. (The BAWE sub-corpus)

Discussion

Hedges are lexical/strategic devices that are frequently used by writers of different genres of texts to present claims with caution. In the present study, they constituted 7.4% of all metadiscourse types in the Iranian Corpus of Learner English (ranked third among all metadiscourse types) and 21% of the total metadiscourse in the BAWE sub-corpus (ranked second among all metadiscourse types). A number of factors are influential concerning the observed underuse of hedges in the Iranian Corpus of Learner English. Among them, previous instruction, culture and L1, and increasing proficiency are the most significant ones.

Role of Previous Instruction

Concerning the role of previous instruction, the most frequently used hedging devices in the Iranian academic, expository essays were modal auxiliaries (e.g., *may*, *should*, *might*, etc.) and epistemic adverbs (e.g., *about*, *almost*, *sometimes*, etc.), which are directly instructed in

courses such as grammar, paragraph development, and essay writing in the Iranian universities. However, other categories including epistemic lexical verbs, which are not instructed explicitly as such, were used less frequently (e.g., *appear* with 9, *argue* with 24, and *assume* with 16 cases of occurrence among all metadiscourse items).

This claim on the role of previous instruction is in accordance with the influential effects of instruction on the use of hedges reported by Jalilifar, Shooshtari, and Mutaqid (2011). They found empirical support for the facilitative effect of explicit instruction in recognizing hedging devices that improved their subjects' language proficiency and reading comprehension scores. In the same vein, Skelton (1988) claims that through the purposeful and explicit teaching of hedges with the aid of a) sensitization exercises, b) rewriting exercises, and c) sets of potential comments, EFL learners will be able to communicate much more successfully with their native counterparts.

Role of Culture and L1

The influence of culture and L1 on the use of hedges has been frequently reported in the literature of research. It is believed that in languages such as Persian, a more direct style of writing (requiring less hedges) is common, while in other languages such as English a more cautious and indirect style (requiring more hedges) is preferred (Hinds, 1987; Hu and Cao, 2011; Mauranen, 1993b). With regard to the differences in using hedges between Persian and English, Abdollahzadeh (2011), for example, asserts that Iranian postgraduates lack full access to a complete repertoire of hedges in comparison with their English counterparts, although the genre under investigation in Abdollahzadeh's paper differs from the current study, and as is known, genre itself has a significant effect on the frequency of use of hedges. Falahati (cited in Rasti, 2011), too, asserts that English writers employ hedges more frequently than Iranian writers.

Role of Increasing proficiency

Increasing proficiency, too, can have influential effects on the use of hedges (Hinkel, 2002), meaning that as Iranian EFL learners become more proficient English language users, they may try to use such features in a more native-like way. Providing proof to this claim, Hyland and Milton (1997) suggest that more proficient EFL learners approximate more closely to native speaker patterns in their use of hedges. In the same vein, Jalilifar, Shooshtari, and Mutaqid (2011) show that EFL learners who have a higher level of language proficiency take better advantage of the explicit instruction of hedging, to the degree that there is an interactive relation between the proficiency level of learners and their metadiscourse knowledge of hedges.

The findings of this study are also in accordance with Azarbad and Ghahraman (2018) and Davoodifard (2006) who found significant differences between the frequencies of hedges in English and Persian abstracts, with underuse of hedges in the Persian ones. They, besides Jalilifar et al. (2011), Nikroo (2010), and Tahririan and Shahzamani (2009), pointed to a number of influential factors in this regard, namely differences between the structure of English and Persian, differences between the culture of Iranian and English writers, language proficiency, and even differences between gender. Differences between Iranian and English natives are also mentioned in studies on genres such as scientific research articles (e.g., Rezanejad, Lari, & Mosalli, 2015), English and Persian editorials (e.g., Tahririan & Shahzamani, 2009), and so forth.

Conclusions

What connects this study to the research literature in learner corpora and hedges is its source of data, that is, EFL learners in university settings. This corpus-based study on hedges

revealed statistically significant differences between Iranian EFL learners' and English-natives' use of hedges. Such differences reside in a number of underlying factors including the fact that English-native writers in the BAWE sub-corpus aimed to show their knowledge of the essay topics, while Iranian EFL writers in the Iranian Corpus of Learner English primarily aimed to show their skills in the English language. Cultural and L1 differences between Persian and native writers as well as the influence of genre, familiarity or lack of awareness concerning hedges resulting from previous instruction, and proficiency level of Iranian EFL learners from whom the essays have been collected in this study, were also key elements that brought about differences with regard to the use of hedges between the two groups of writers.

Generally, it is assumed that direct and explicit instruction of hedges according to the norms of the genre of interest seems to be the best strategy in order to alleviate the overuse/underuse of hedges among Iranian EFL learners in university settings. For this, applying corpus software and online websites could be a great help. Using software such as AntConc through DDL practices can result in great changes in the Iranian university-level learners' production of hedges in their academic essays. For this to happen, teachers should be encouraged to learn the function of hedges first, and then through becoming familiar with corpora and corpus software and integrative approaches to computer-assisted language learning (CALL) (Warschauer, 1996) such as flip teaching (Bergmann & Sams, 2012), they can provide learners with enough sources to learn for themselves and reach learner autonomy in the use of hedges.

References

Abdi, M., & Behnam, B. (2014). Cross-cultural analysis of hedges in English and Persian medical journals. *International Journal of Language Learning and Applied Linguistics World, 7*(1), 92-107.

Abdollahzadeh, E. (2011). Poring over the findings: Interpersonal authorial engagement in applied linguistics papers. *Journal of Pragmatics, 43*(1), 288-297. https:/ /doi.org /10.1016/ j.pragma.2010.07.019

Ackermann, K., Biber, D., & Gray, B. (2011, July 20-22). An academic collocation list [Paper presentation]. Corpus Linguistics, Birmingham, UK.

Ädel, A. (2006). *Metadiscourse in L1 and L2 English*. Amsterdam, The Netherlands: John Benjamins Publishing Company.

Ädel, A. (2010). Just to give you kind of a map of where we are going: A taxonomy of metadiscourse in spoken and written academic English. *Nordic Journal of English Studies, 9*(2), 69-97.

Ahn, M., & Yap, F. H. (2015). Evidentiality in interaction: A pragmatic analysis of Korean hearsay evidential markers. *International Journal sponsored by the "Foundations of Language", 39*(1), 46-84. https://doi.org/10.1075/sl.39.1.03ahn

Anthony, L. (2017). AntFileConverter (Version 1.2.1) [Computer software]. Waseda University. https://www.laurenceanthony.net/software/antfile converter/

Atkins, S., Clear, J., & Ostler, N. (1992). Corpus design criteria. *Literary and Linguistic Computing, 7*(1), 1-16. https://doi.org/10.1093/llc/7.1.1

Azarbad, E., & Ghahraman, V. (2018). A comparative study on the English to Persian translation of hedges in the abstracts of MA theses in English translation studies. *Journal of Language and Translation, 8*(3), 57-67.

Barlow, M. (2017). MonoConc Pro-Semester (version 2.2.) [Computer software]. Althestan. http://www.michaelbarlow.com

Bergmann, J., & Sams, A. (2012). *Flip your classroom: Reach every student in every class every day*. Eugene: International Society for Technology in Education.

Bouhlal, F., Horst, M., & Martini, J. (2018). Modality in ESL textbooks: Insights from a contrastive corpus-based analysis. *Canadian Modern Language Review, 74*(2), 227-252. https://doi.org/10.3138/cmlr.3075

Burrough-Boenisch, J. (2005). NS and NNS scientists' amendments of Dutch scientific English and their impact on hedging. *English for Specific Purposes, 24*(1), 25-39. https://doi.org/10.1016/j.esp.2003.09.004

Callies, M. (2015). Learner corpus design and methodology. (2015). In S. Granger, G. Gilquin & F. Meunier (Eds.), *The Cambridge handbook of learner corpus research* (pp. 7-182). Cambridge: Cambridge University Press.

Channell, J. (1994). *Vague Language*. Oxford: Oxford University Press.

Cobb, T. (2016). Corpus Builder (v.2.3).http://www.lextutor.ca/cgi-bin/tools /corp_build /index. pl?uppers=50

Davoodifard, M. (2006). A contrastive analysis of hedging in English and Persian research articles: Linguistic and cultural variations across languages and disciplines [Unpublished M.A. thesis].University of Esfahan.

Dontcheva-Navratilova, O. (2016). Cross-cultural variation in the use of hedges and boosters in academic discourse. *Prague Journal of English Studies, 5*(1), 163-184. https: //doi.org /10.1515 /pjes-2016-0009

Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford: Oxford University Press.

Ellis, R., & Barkhuizen, G. P. (2005). *Analyzing learner language*. Oxford: Oxford University Press.

Granger, S. (2003). The international corpus of learner English: A new resource for foreign language learning and teaching and second language acquisition research. *TESOL Quarterly, 37*(3), 538-546. https://doi.org/10.2307/3588404

Granger, S. (2012). How to use foreign and second language learner corpora. In A. Mackey & S.M. Gass (Eds.), *Research methods in second language acquisition: A practical guide* (pp. 7–29). London: Blackwell Publishing.

Hawkey, R. & Barker, F. (2004). Developing a common scale for the assessment of writing. *Assessing Writing, 9*(2), 122–159. https://doi.org/10.1016/j.asw.2004.06.001

Hawkins, J.A., & Filipović, L. (2012). *Criterial features in L2 English: Specifying the reference levels of the Common European Framework*. Cambridge: Cambridge University Press.

Hinds, J. (1987). Reader versus writer responsibility: A new typology. In U. Connor & R. B. Kaplan (Eds.), *Writing across languages: Analysis of L2 text* (pp. 141-15). Reading, MA: Addison Wesley.

Hinkel, E. (2002). *Second language writers' text*. Mahwah, NJ: Lawrence Erlbaum.

Hu, G., & Cao, F. (2011). Hedging and boosting in abstracts of applied linguistics articles: A comparative study of English-and Chinese-medium journals. *Journal of Pragmatics, 43*(11), 2795-2809. https://doi.org/10. 1016/j.pragma.2011.04.007

Hunston, S. (2002). *Corpora in applied linguistics.* Cambridge: Cambridge University Press.

Hyland, K. (1996). Talking to the academy: Forms of hedging in science research articles. *Written Communication, 13*(2), 251-81.

Hyland, K. (1998). *Hedging in scientific research articles*. Amsterdam, The Netherlands: John Benjamins Publishing Company.

Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. London: Continuum.

Hyland, K. (2017). Metadiscourse: What is it and where is it going? *Journal of Pragmatics, 113*(1), 16-29. http://dx.doi.org/10.1016/j.pragma. 2017.03.007

Hyland, K., & Milton, J. (1997). Qualification and certainty in L1 and L2 students' writing. *Journal of Second Language Writing, 6*(2), 183-205. https://doi.org/10.1016/S1060-3743(97)90033-3

Jalilifar, A., Shooshtari, Z. G., & Mutaqid, S. (2011). The effect of hedging instruction on reading comprehension for Iranian university students. *Research in Applied Linguistics, 2*(1), 69-89.

Larsson, T. (2017). A functional classification of the introductory it pattern: Investigating academic writing by non-native-speaker and native-speaker students. *English for Specific Purposes, 48*, 57-70. https://doi.org/10.1016/j.esp.2017.06.001

Li, Q., & Pramoolsook, I. (2015). Research article abstracts in two subdisciplines of business-move structure and hedging between management and marketing. *English Language Teaching, 8*(1), 52-62.

Lozano, C., & Mendikoetxea, A. (2013). Learner corpora and second language acquisition: The design and collection of CEDEL2. In N. Ballier, A. Díaz-Negrillo & P. Thompson (Eds.), *Automatic treatment and analysis of learner corpus data* (pp. 65–100). Amsterdam, The Netherlands: John Benjamins Publishing Company.

Markkanen, R., & Schröder, H. (1997). Hedging: A challenge for pragmatics and discourse analysis. In R. Markkanen & H. Schröder (Eds.), *Hedging and discourse: Approaches to the analysis of a pragmatic phenomenon in academic texts* (pp. 3-18). Berlin: Walter de Gruyter.

Martikainen, H. (2018). Modal markers as potential sources of distortion intranslated medical abstracts. *Educational Sciences: Theory and Practice, 18*(4), 917-934.

Mauranen, A. (1993a). Contrastive ESP rhetoric: Metatext in Finnish-English economics texts. *English for Specific Purposes, 12*(1), 3-22. https://doi.org/10.1016/0889-4906(93)90024-I

Mauranen, A. (1993b). *Cultural differences in academic rhetoric: A textlinguistic study.* Frankfurt: Peter Lang.

Mauranen, A. (1997). Hedging in language revisers' hands. In R. Markkanen & H. Schröder (Eds.), *Hedging and discourse: Approaches to the analysis of a pragmatic phenomenon in academic texts* (pp. 115-133). Berlin: Walter de Gruyter.

Mauranen, A. (2010). Discourse reflexivity-a discourse universal? The case of ELF. *Nordic Journal of English Studies, 9*(2), 13-40.

Nikroo, P. (2010). The relationship between the knowledge of sentence structure and quality of translation. *International Journal of Language Studies, 4*(1), 1-58.

Plappert, G. (2019). Not hedging but implying: Identifying epistemic implicature through a corpus-driven approach to scientific discourse. *Journal of Pragmatics, 139*, 163-174. https://doi.org/10.10 16/j.pragma.2018.09.001

Rasti, I. (2011). Involving the reader in the text: Engagement markers in native and non-native student argumentative essays [Unpublished doctoral dissertation]. University of Liverpool, UK.

Rayson, P. (2019). Log-likelihood and effect size calculator [Computer program]. http:// ucrel.lancs.ac.uk/llwizard.html

Resche, C. (2015). Hedging in the discourse of central banks. *Studies in Communication Sciences, 15*(1), 83-92. https://doi.org/10.1016 /j.scoms.2014.12.008

Rezanejad, A., Lari, Z., & Mosalli, Z. (2015). A cross-cultural analysis of the use of hedging devices in scientific research articles. *Journal of Language Teaching and Research, 6*(6), 1384-1392.

Sabzevar, A., Haghverdi, H., & Biriya, R. (2020). A corpus-based analysis of epistemic stance adverbs in essays written by native English speakers and Iranian EFL learners. *Journal of English Language Pedagogy and Practice, 12*(25), 32-52. https://doi.org/10.30495/JAL.2020.674922

Skelton, J. (1988). The care and maintenance of hedges. *ELT Journal, 42*(1), 37-43. https://doi.org/10.1093/elt/42.1.37

Sun, X., & Hu, G. (2020). Direct and indirect data-driven learning: An experimental study of hedging in an EFL writing class. *Language Teaching Research, 24*(5), 1-29. https://doi.org/10.1177/ 1362168820954459

Tahririan, M. N. & Shahzamani, M. (2009). Hedging in English and Persian editorials: A contrastive study. *Iranian Journal of Applied Linguistics, 12*(1), 199-221.

Tham, D. (2013). *Expository eureka: Model expository essays for today's secondary school students*. Singapore: Marshall Cavendish Editions.

Tognini-Bonelli, T. (2001). *Corpus linguistics at work.* Amsterdam, The Netherlands: John Benjamins Publishing Company.

Tsukamoto, S. (2002). KWIC concordance for windows: Easy access to corpora. In T. Saito, J. Nakamura & S. Yamazaki (Eds.), *Language and computers, English corpus linguistics in Japan* (pp. 327–340). Amsterdam: Rodopi.

Vande Kopple, W. J. (1985). Some exploratory discourse on metadiscourse. *College Composition and Communication, 36*(1), 82-93. https://doi.org/10.2307/357609

Varma, S. B., & Tan, H. (2015). Hedges used by ESL students in single-sex and mixed-sex informal conversations. *Language & Communication, 2*(2), 169-182.

Vass, H. (2017). Lexical verb hedging in legal discourse: The case of law journal articles and supreme court majority and dissenting opinions. *English for Specific Purposes, 48*, 17-31. https://doi.org/10.1016/ j.esp.2017.07.001

Warschauer, M. (1996). Computer-assisted language learning: An introduction. In S. Fotos (Ed.), *Multimedia language teaching* (pp. 3-20). Tokyo: Logos International.

Weisi, H., & Asakereh, A. (2020). Hedging devices in applied linguistics research papers: Do gender and nativeness matter? *Glottotheory*. Advance online publication. https://doi.org/10.1515/glot-2020-2013

Wulff, S., & Gries, S. T. (2011). Corpus-driven methods for assessing accuracy in learner production. In P. Robinson, (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 61-87). Amsterdam, The Netherlands: John Benjamins Publishing.

Appendices

Appendix A: Raw Frequency of Hedges in the Iranian Corpus of Learner English and the BAWE Sub-corpus

Items	ICLE	the BAWE sub-corpus	Items	ICLE	the BAWE sub-corpus
about	167	5	almost	146	5
apparent	4	4	apparently	10	3
appear	9	26	appeared	2	2

appears	3	26	approximately	16	7
Argue	15	0	argued	5	4
Argues	4	0	around	24	4
Assume	16	3	assumed	3	7
broadly	3	0	certain amount	1	1
certain extent	0	1	certain level	4	1
claim	27	9	claimed	6	20
Claims	0	10	could	156	68
couldn't	4	3	doubt	2	0
doubtful	0	1	essentially	4	0
estimate	7	1	estimated	5	1
fairly	2	4	feel	6	22
feels	3	4	felt	0	8
frequently	1	26	from my perspective	0	0
from our perspective	0	0	from this perspective	0	0
generally	76	38	guess	4	1
indicate	15	18	indicated	3	8
indicates	11	9	in general	18	5
in most cases	9	1	in most instances	0	0
in my opinion	0	0	in my view	8	2
in this view	0	0	in our opinion	0	0
in our view	0	0	largely	4	6
likely	100	58	mainly	21	24
may	246	162	maybe	76	4
might	176	45	mostly	52	2
often	126	127	on the whole	2	0
ought	1	1	perhaps	48	26
plausible	0	0	plausibly	0	0
possible	120	22	possibly	19	12
postulate	0	0	postulated	0	1
postulates	0	0	presumable	0	0
presumably	0	0	probable	0	2
probably	42	10	quite	54	11
rather x	12	0	relatively	7	20
roughly	1	3	seems	76	34
should	340	12	sometimes	148	17
somewhat	13	5	suggest	4	37
suggested	7	51	suggests	10	61
suppose	3	1	supposed	14	4
supposes	0	1	suspect	0	0

suspects	0	0	tend to	48	42
tended to	2	8	tends to	13	8
to my knowledge	0	0	typical	21	19
typically	19	9	uncertain	0	3
uncertainly	0	0	unclear	1	6
unclearly	0	0	unlikely	5	7
usually	120	14	would	172	84
wouldn't	41	11	***TOTAL***	***2964***	***1327***

Note: ICLE stands for the Iranian Corpus of Learner English