Trends of Corpus Linguistics Used in English for Specific Purposes Research: A Case of Asian ESP Journal

Jenjira Jitpaiboon[1] and Atichat Rungswang[1*]

[1]*Faculty of Liberal Arts, King Mongkut's Institute of Technology Ladkrabang*
**atichat.ru@kmitl.ac.th*

ABSTRACT

This corpus-based study aims to investigate the trends of corpus linguistics used as a methodology in the English for Specific Purposes (ESP) research based on corpora of research article (RA) abstracts and conclusions published in the Asian ESP journal between 2005 and 2020. The findings reveal that 52 research articles out of 452, accounting for 11.5%, adopted corpus linguistics as a methodology. The number seems to be highest between 2011 and 2015 and tends to fall dramatically between 2016 and 2020. In terms of research approach, the corpus-based approach appears to outnumber the corpus-driven overwhelmingly. Swales' CARS model and study on research articles were strongly connected with corpus linguistics. However, the pedagogical application of corpus linguistics in the ESP domain appears to be underresearched. This study demonstrates how corpus linguistics can be applied as a research tool in trends investigation and contributes to the area of both corpus linguistics and ESP research.

Keywords: Corpus Linguistics; Corpus-based approach, Corpus-driven approach, trend analysis, English for Specific Purposes (ESP)

Introduction

According to McCarthy and O'Keeffe (2010), *corpus linguistics* becomes the common term known by linguists referring to the act of "searching through screen after screen of concordance lines and wordlists generated by computer software, in an attempt to make sense of phenomena in big texts or big collections of smaller texts" (p.3). A term *corpus* itself is defined as a body of texts collected from spoken or written sources which can be particularly of use to conduct a large-scale investigation of words in context, identify specialized vocabulary for more understanding, and build word lists for classroom use and autonomous study (Coxhead, 2013).

The applications of corpus linguistics are diverse. It has been applied to the empirical analysis of language and adopted in "language teaching and learning, discourse analysis, literary stylistics, forensic linguistics, pragmatics, speech technology, sociolinguistics, and health communication" (McCarthy & O'Keeffe, 2010, p.7). Using software tools in corpus linguistics facilitates the analysis of many searches. It offers "invaluable statistical information about co-occurrences, trends, tendencies, frequency, and distributions", making it a faster method than any other manual (Vo & Carter, 2010, p.305).

Two main approaches are applied in corpus linguistics research, that is, corpus-based and corpus-driven. According to Vo and Carter (2010), in the corpus-based approach, intuitions are placed first, and corpora are employed to test and validate them as sources of empirical data. On the other hand, in a corpus-driven approach, "corpora themselves are the data from which creative language uses are uncovered" (p.310). In other words, Tognini Bonelli (2001) defines the corpus-driven approach in the way that "observation leads to hypothesis leads to generalization leads to unification in a theoretical statement" (p.85), while the corpus-based approach is based on previously held beliefs or pre-existing rules.

English for Specific Purposes (ESP) research has also been revolutionized by the corpus linguistics approach, which has made it possible and convenient to investigate large sets of texts and has contributed to the generalizability of findings (Salmani-Nodoushan, 2020). In light of this, Lesiak-Bielawska (2015) mentions that corpus linguistics seems to have a lot to offer to English for ESP practitioners; for example, they can make use of mega-databanks which allows them to access authentic written and spoken discourse, and with the help of concordancing software, they can also develop specialized corpora, compare learner and expert texts, as well as examining the distribution of specialized grammatical and lexical features. Lesiak-Bielawska suggests that corpora have significantly provided numerous alternatives available to ESP practitioners and have raised the significance of research activity in the new forms.

As a result, the study aimed to examine the trends of corpus linguistics, including the extent and areas used in the ESP research context. The findings will significantly benefit graduate students, researchers, professors, and policymakers to follow the trends or fill the gap in the underrepresented areas. Furthermore, the study also helped demonstrate corpus linguistics's use in the trends investigating process.

Literature Review

Friginal and Hardy (2013) note that nearly all published research articles in the subfields of applied linguistics, such as ESP and English for Occupational Purposes (EOP), have highly become corpus-based, taking advantage of some forms of corpus analysis; nevertheless, corpus-based studies still are underresearched, particularly in terms of spoken discourse analysis. ESP and EOP studies on workplace discourses often focus on practical and pragmatic applications of communication patterns.

A corpus has also had an enormous impact in English for Academic Purposes (EAP) both in terms of direct (as a pedagogic tool) and indirect (as a reference and evidence) ways (Charles, 2013). Based on Charles's comments, corpora have contributed much to EAP in indirect use, especially in written discourse. Corpus-based approaches have become the norm in EAP, e.g., to describe specific features of academic discourse and its phraseology, reveal the characteristics of different disciplines and genres, examine recurrent word sequences, compile academic wordlists, and identify signals of moves and move boundaries. As for the direct application of corpus data with students, Charles identifies that research has recently also begun to focus on evaluating the effectiveness of direct corpus approaches and their acceptability to students, but there is a scarcity of work on the pedagogical applications of genre analysis.

The dominance of the corpus linguistics approach in written discourse is also pointed out by Johns (2013), who elaborates with solid evidence found in research studies submitted to the *English for Specific Purposes* journal (ESPj) and other research scenes. Similarly, Gollin-Kies (2014) conducted a study to compare research articles in two journals: *English for Specific Purposes journal* (ESPj) and the *Journal of English for Academic Purposes* (JEAP) published between 2003 and 2012 to examine the research methods most used and the research paradigms underrepresented in the field. The findings revealed that qualitative analysis of corpus based on written discourse was overwhelmingly found in both journals about frequency and distribution of linguistic or rhetorical features. Sa-ngiamwibool (2014) examined 70 articles in the *Asian ESP Journal* (AESPJ) published between 2011 and 2013, and the results of the study revealed that genre analysis and corpus linguistics were prevalent. The topics studied during this period primarily included lexical bundles in journalistic discourse and research article abstracts. Sa-ngiamwibool projected that corpus linguistics tended to continue its influence over the next decade. Salmani-Nodoushan (2020) notes that in the past two decades, most ESP practitioners and teachers have changed their research focus from theoretical issues to methodological and analytical studies on topics such as genres, corpora, and metadiscourse. According to Salmani-Nodoushan's analysis, corpus linguistics/analysis, discourse analysis, and computer and digital technology will continue to influence ESP research.

Methodology

While other research has already investigated the trends of general research methods in ESP-related journals using a manual approach, this current study concentrates explicitly on the trends

of corpus linguistics used as a methodology in AESPJ, based on the findings from the corpus analysis tool.

Most articles published in AESPJ are from Asia, Africa, and the Middle East. In addition, the journal's name indicates that Asia is the main target of this journal. Since both researchers in this study are in Thailand, it would be highly advantageous for Thai and other Asian researchers to observe the trends of international journals like AESPJ, which are the primary target.

This research adopts a corpus-based approach applying the previous research findings as a guideline in the investigation. The period of research publication was also expanded to cover the research articles published between 2005 and 2020, accounting for 16 years.

The research data was based on corpora of 446 research article (RA) abstracts and six conclusions published in the AESPJ between 2005-2020. Most texts in the corpora (N=446) were from RA abstracts which, as Sánchez (2018) suggested, are typically required by most journals to be included for publication. RA abstracts are considered a significant part of the RAs, acting as "an indicator of whether the research idea has been well established and disseminated" (Fatma & YAĞIZ, 2020, p.391). Most importantly, abstracts "offer preliminary notions about the research" (Sánchez, 2018, p.215). Hengl and Gould (2002) note that the "abstract should be short but give the overall idea: what was done, what was found and what are the main conclusions" (p.1). Despite its significance, abstracts were found missing in six research articles; therefore, the researcher decided to adopt the RA conclusion as the substitute, given that the RA conclusion's primary purpose is "to summarize the research by highlighting the findings, evaluating and pointing out possible lines of future research as well as suggesting implications for teaching and learning" (Ruiying & Allison, 2003, p.380). RA abstracts and conclusions represent the texts corresponding to the purpose of this study, that is, to examine the trends in research ideas and research approaches.

All available sixteen volumes of the AESPJ, published between 2005 and 2020, were retrieved from its official website. They were sifted through to extract only research articles (RAs). Forwards, company reviews, book reviews, and business articles were excluded. After that, 446 RA abstracts and six conclusions were then divided into three timeframes: 2005-2010 (6 years), 2011-2015 (5 years), and 2016-2020 (5 years). There were only six RA abstracts in 2005; hence they were included in the first period.

All RA abstracts were then transformed into plain text or *.txt* format, and a freeware *AntConc* (Anthony, 2004) was employed for data analysis. The search terms corpus and corpora were used at the concordance function to extract only research articles employing corpus linguistics as a methodology. Then the researchers checked the file names displayed in the next column, next to the concordance lines. All articles in which search terms corpus and corpora were found were then separated into three folders and three timeframes. Fifty-two articles from the selection process were then run again for trend analysis.

Results

Overall results

Based on the corpus data analysis, out of 446 RA abstracts and six conclusions published between 2005 and 2020 in the AESPJ, there are 52 research articles in which the search terms *corpus* and *corpora* were found, accounting for 11.5% when compared with the total number of 452 articles.

Table 1 The number of ESP research applying corpus linguistics as a methodology in the AESPJ between 2005 and 2020

Years	Total no. of articles	No. of articles with the terms *corpus* and *corpora*	Percentage of corpus articles compared with total no. of articles
2005-2010	58	9	15.5%
2011-2015	97	22	22.7%
2016-2020	297	21	7.1%
Total	**452**	**52**	**11.5%**

From Table 1, when divided into three timeframes, the number of articles that applied corpus linguistics as a methodology seems to be the highest (22.7%) between 2011 and 2015. It appears to be the least between 2016 and 2020.

The results correspond with Sa-ngiamwibool (2014) 's findings that corpus linguistics used in AESPJ articles was prevalent between 2011 and 2013. However, the trends seem to decline between 2016 and 2020, which may be concluded that Sa-ngiamwibool's (2014) projection of corpus linguistics' rising trends in the following decades seems incorrect.

Corpus-based vs. Corpus-driven

As discussed earlier, corpus-based and corpus-driven are the two main approaches adopted in corpus linguistics research. The corpus-based approach checks the researcher's intuition, while the corpus-driven approach is adopted with an open mind to see what patterns emerge (Tognini – Bonelli, 2001). When the terms *corpus-based, corpus-based, corpus-driven, and corpus-driven were searched in the corpus of 52 RA abstracts and conclusions, it appears that the corpus-based and corpus-based terms* overwhelmingly outnumber *corpus-driven* and *corpus-driven,* as presented in Table 2.

Table 2 Frequency corpus-based and corpus-driven in AESPJ between 2005 and 2020

Years	Corpus-based/Corpus base	Corpus-driven/Corpus driven
2005-2010	1	0
2011-2015	6	0
2016-2020	3	2
Total	**10**	**2**

Trends of corpus studies in AESPJ between 2005 and 2010

Based on Table 1, there are only 9 RA abstracts and conclusions in which corpus and corpora search terms were found during this period. The corpus contains 1,555 tokens with 559-word lists. McCarthy and O'Keeffe (2010) suggest that "small, carefully targeted corpora (by which we commonly mean corpora of fewer than a million words of running text) have proved to be a powerful tool for the investigation of special uses of language, where the linguist can 'drill down into the data in immense detail using a full armory of software and shed light on particular uses of language" (p.6). The corpus of these 9 RA abstracts and conclusions in this study has also been purposefully selected and was faithfully based on the corpus analysis.

The trends investigation in this study was grounded on top ten-word lists, and their collocates run on *AntConc* based on their frequency, as demonstrated in Table 3. Only key content words and one lemma occurring first in the list were selected. Content words refer to nouns, verbs, adjectives, and adverbs (Scott & Tribble, 2006). Function words or grammatical information were excluded since they do not carry meaning (Bortolato, 2016). Concordance lines were also rechecked for better comprehension. The same criteria were applied to selecting and analyzing word lists and collocates. The first word, *English*, emphasized the representativeness in which the researcher aimed to examine the trends of corpus linguistics in the area of *English* for specific purposes.

Table 3 Ten most frequent word lists and their ten most frequent collocates between 2005 and 2010

No	Word Lists	Collocates									
1	English	RAs	Native	Language	Business	Writers	Speaking	Published	Persian	Medical	Literature
2	Study	Present	Examines	Uses	Suggests	Reveal	Observational	Incorporated	Corpus	Consists	Case
3	Research	Articles	Writing	Medical	Tried	Texts	Published	Persian	Linguistics	Forty	English
4	Corpus	Data	Used	Study	Showed	Revealed	Relevant	Pilot	Parallel	Current	Based
5	Articles	Research	ELT	Selected	RAs	Experimental	Comprised	-	-	-	-
6	Language	English	Teaching	Student	Real	Figurative	comprehension	challenges	-	-	-
7	Journals	International	TESL	Swales	Persian	Medical	Local	Indicating	English	ELT	
8	Persian	Research	Professional	Made	Journals	English	EFL	Capitalized	-	-	-
9	International	Journals	Trade	Scale	English	Counterparts	Communications	-	-	-	-
10	Results	Support	Show	Practices	Obtain	Effective	Directly	Confirmed	-	-	-

The corpus linguistics approach applied in AESPJ during this timeframe may focus on analyzing *RAs* and native and *medical* discourse. Written texts may be an emphasis in this period since the word *speaking*, when checked with concordance, mainly refers to *native English speaking*, not spoken texts. The *Persian* language seems to be popular to be investigated together with the English language in the AESPJ in this period, as seen in Figure 1. *Literature* is another area that tends to be explored in this phase.

the use of metadiscourse elements in Persian and English research articles. The research
. The intralingual analysis showed that both Persian and English used interactive resources more
the academic genre. Compared with English, Persian capitalized on more interactive resources, which
English RAs written in English by Persian EFL writers rejected by international journals
of Persian RAs published in professional Persian journals of the same field in
to find out how English and Persian made use of metadiscourse elements, and
more interactive resources, which shows that Persian puts a premium on textuality at
compared with a parallel corpus of Persian RAs published in professional Persian journals
rentiating factors between published English and Persian research articles at the level of
in the two languages, English and Persian. The intralingual analysis showed that both

Figure 1 Concordance lines of 'Persian'

There is one *observational* study conducted on linguistic metaphors. The third most frequent word found in this corpus, *research*, still emphasizes the dominance of RAs in corpus studies in ESP in this period (See Figure 2). Pedagogical application is noticed from the collocates *teaching*, *ELT* (English Language Teaching), *TESL* (Teaching English as a Second Language), and *EFL* (English as a Foreign Language). Swales' CARS model is revealed in the corpus.

, allowing them to better understand published research articles and facilitating the process of
factors between published English and Persian research articles at the level of move
and facilitating the process of writing research articles for publication.
the discussion sections of applied linguistics research articles (RAs) from the perspective of
based study uses two corpora of 40 research articles selected from two TESL journals
metadiscourse elements in Persian and English research articles. The research tried to find
of the discussion sections of medical research articles. We obse group consisted of

Figure 2 Concordance lines of 'Research'

Trends of corpus studies in AESPJ between 2011 and 2015

Similar to results revealed in the first period, *English* is the first most repeated word in 22 RA abstracts and conclusions published between 2011 and 2015. *English* is strongly collocated with *American, Philippines, Native, Chinese Speakers, Proficiency, Abstracts, Vocabulary, University,* and *Tests*. According to the concordance, three varieties of English, namely American, Philippines, and Chinese, seem to be often investigated. RA *abstracts* and *vocabulary* may still be the research focus of corpus studies. *University* settings may be among the areas of research interest.

Table 4: Ten most frequent word lists and their ten most frequent collocates between 2011 and 2015

No	Word Lists	Collocates									
1	English	American	Philippines	Native	Chinese	Speakers	Proficiency	Abstracts	Vocabulary	University	Tests
2	Corpus	Based	Word	RA	Linguistic	Incorporating	Data	Containing	Annotated	Various	Tools
3	Academic	Writing	Words	Purposes	Register	World	Vocabulary	Various	Use	Typical	Texts
4	Research	Article	Business	Previous	Mathematics	Writing	Three	Studies	Space	Questions	Quantitative
5	Study	Examines	Aimed	Suggested	Shows	Reveal	Provided	Profiles	Present	Pilot	Lexical
6	Business	Research	Meetings	Word	Twenty	Specific	Purposes	Subjects	Students	Science	Register
7	Lexical	bundles	Verbs	Coverage	Word	Items	Competence	Vocabulary	Study	Students	Specific
8	Word	List	Corpus	Families	Academic	Million	Lexical	Four	Business	Top	Token
9	Vocabulary	Technical	Profiles	Learning	Active	Students	Specific	Specialized	Selected	participants	Online
10	Writing	Academic	Mathematics	Teaching	Student	Specific	Resemble	Research	Practices	Letter	Instructors

Research may appear to be highly corpus-*based* in this period (See Figure 3), and most of the corpora tend to be *annotated*.

Figure 3 Concordance lines of 'Based'

As shown in Figure 4, academic *writing* seems to be highly interesting to researchers. This matches the findings in ESPj and JEAP previously conducted by Gollin-Kies (2014).

Figure 4 Concordance lines of 'Writing'

Writing, Words, and *Purposes* seem to be associated with English for academic purposes (EAP). In addition, this period also focuses on academic *register* and *vocabulary* and is strongly linked with the analysis of the research *article* (See Figure 5). Hence, corpora seem to contribute much to EAP, per Charles' (2013) advice.

Figure 5 Concordance lines of 'Article'

In addition, *Lexical bundles* are primarily found in this corpus (See Figure 6). It proves Sangiamwibool's (2014) findings indicated that the topics between 2011 and 2013 mainly involved lexical bundles. It also confirms the interest of ESP researchers in the distribution of specialized grammatical and lexical features to study the authentic use in the actual context, which can be facilitated by using corpus linguistics as a methodology, as Lesiak-Bielawska (2015).

terms of the frequency and function of lexical bundles. A corpus of 200 research article abstracts
istribution of functional and structural types of lexical bundles and their probable relations as used
cture could possibly be realized through specific lexical bundles and vocabulary. This study examines vocabulary
unctions were described. The findings reveal that lexical bundles employed in journalistic materials are mostly
of the bundles also shows that the lexical bundles found in the Iranian corpus included
paper aims to investigate the use of lexical bundles in newspapers. It aims to shed
cific move structures, move-signaling words, and lexical bundles in RAs provide valuable information for
This study examines vocabulary use, particularly lexical bundles, in the introductions of computer science
reveal that Iranian authors use more 4-word lexical bundles in their writing compared to their
five-word, four-word, and three-word lexical bundles that reflect the specific rhetorical functions
assigned to these frequently used expressions is lexical bundles. This paper aims to investigate the
Daily were selected to form a corpus. Lexical bundles were identified with the use of
major moves, move-signaling words and meaningful lexical bundles were identified. Pedagogically, the corpus-based

Figure 6 Concordance lines of 'Bundles'

Trends of corpus studies in AESPJ between 2016 and 2020

In contrast with the findings in the previous two periods, *study* is the most repeated word in 21 RA abstracts and conclusions published between 2016 and 2020, which may be regarded as the norm of RA abstract writing. *English* is strongly associated with *language, native, relative, writers, variable, used, Turkish, teaching, spoken,* and *speakers*. There are two corpus-*driven* research articles; however, the researchers still favor corpus-*based* in this period. The *academic* area tends to focus on research implementation, particularly in the *academic promotional* genre. *The qualitative* corpus approach appears to be often employed.

Table 5: Ten most frequent word lists and their ten most frequent collocates between 2016 and 2020

No	Word Lists	Collocates									
1	**Study**	Present	Examines	Concludes	Aimed	Used	Tried	Seminal	Seeks	Provides	Models
2	**English**	Language	Native	Relative	Writers	Variable	Used	Turkish	Teaching	Spoken	Speakers
3	**Corpus**	Based	VNMC	Driven	Compiled	Word	Used	Thesis	Reveals	Qualitative	Lecture
4	**Used**	Frequently	First	Veterinary	Turkish	Study	RVs	Rarely	Qualitative	Predominantly	Nominalization
5	**Research**	Articles	Design	Terms	Related	Published	Publication	Process	Past	Little	Investigating
6	**Academic**	Writing	Promotional	Words	Purposes	Genres	Discourse	Word	Teaching	Spoken	Research
7	**Rhetorical**	Moves	Repertoire	Students	Questions	Occurring	Novice	Involves	Features	Examines	Eight
8	**Students**	Graduate	Novice	Journalism	Writing	Vietnamese	Undergraduate	Rhetorical	Postgraduate	Native	Made
9	**Words**	List	Including	Academic	University	Synonym	Running	Related	Pakistani	Occurring	Hold
10	**Analysis**	Genre	Collocational	Unveils	Thematic	Textual	Shows	Score	Resulted	Qualitative	Move

reporting the results along with their rhetorical configurations. This study discloses the impact
The rhetorical construction of discourse is constantly changing
impact of research design on the rhetorical conventions of students' F&D sections.
NNEWs the various functions and the rhetorical effects that RVs can have on
was carried out to explore the rhetorical features, i.e., occurrences, salience, and
showed that the most commonly occurring rhetorical move pattern and stance features for
in academic discourse analysis that involves rhetorical move patterns and metadiscourse stance features.
was an inappropriate usage of both rhetorical move patterns and stance features in
was carried out to analyze the rhetorical moves. Afterward, a bottom-up approach
on genre analysis to investigate the rhetorical moves and communicative functions of the
, the paper sought to identify the rhetorical moves and the functional constituent of
This study examines rhetorical moves and their linguistic realizations including
realizations. The findings demonstrated that the rhetorical moves and their constituent steps of
, verb function, and sentence voice. However, rhetorical moves considerably exemplified a variety of
most common in the corpus. All rhetorical moves have utilized verbs in this
verbal groups in experiential meaning. The rhetorical moves in the twelve (12) brochures were
findings, and non-standard configurations of rhetorical moves, indicating the students' novice rhetorical
This study examines the rhetorical moves of the NNES undergraduate students'
he communicative functions were determined. Eight rhetorical moves were identified and in terms
action or use classroom management or rhetorical questions. However, NES lecturers use more
to mediate the enhancement of students' rhetorical repertoire.
rhetorical moves, indicating the students' novice rhetorical repertoire. Meanwhile, the discrepancies influen

Figure 7 Concordance lines of 'Rhetorical'

Analysis of *rhetorical moves* and *repertoire* tends to be most often conducted (See Figure 7).

Most research in this period may appear to be linked with *graduate* students (See Figure 8) and *novice* researchers. The field of *journalism* tends to be the research focus in this period. Data analysis relates to *genre, collocation, theme, text, qualitative,* and *move.*

experts, EAP teachers, and academic writing students.
designing curriculum and textbooks for journalism students.
by Korean science and engineering graduate students and junior researchers. Relative clauses promote
by Korean science and engineering graduate students and junior researchers differ from those
non-native science and engineering graduate students and junior researchers use English relative
of English so that the journalism students are exposed to various writing styles
past research in that many graduate students avoid the use of first-person
course to enhance lecturers' skills and students' comprehension of lectures.
passive was rarely used. Turkish postgraduate students differed from the other two groups
-person pronouns in their writing. Graduate students, especially those in Asia and in
design on the rhetorical conventions of students' F&D sections. Hence, this paper
in gathering academics, researchers, experts, and students in conferences. Thus, this study tried
, especially for novice researchers and graduate students in medical sciences as well as
(M.A.) theses written by Vietnamese students in terms of their denotative and
rhetorical moves of the NNES undergraduate students' iterative findings and discussion (F&D)
that, English native and non-native students made use of similar lexical choices
postgraduate English for Academic Purposes (EAP) students, novice researchers, conference sponsors, confer
onfigurations of rhetorical moves, indicating the students' novice rhetorical repertoire. Meanwhile, the di
pedagogy to mediate the enhancement of students' rhetorical repertoire.
pronouns much more frequently than doctoral students. The one exception was with TESOL

Figure 8 Concordance lines of 'Students'

Discussion and conclusion

The number of articles with the terms *corpus* and *corpora* seems to be highest between 2011 and 2015 (22.7%). The results correspond with Sa-ngiamwibool (2014) 's findings which indicate that corpus linguistics was prevalent between 2011 and 2013.

However, between 2016 and 2020, despite the high number of research articles published during this period (N = 297), the number of articles in which the search terms *corpus* and *corpora* were found tended to plummet (N = 21, accounted for 7.1%). It seems that ESP researchers might turn their focus to other methods. The results contradict Sa-ngiamwibool (2014) projection of rising corpus linguistics trends in the following decades.

The terms *corpus-based, corpus base,* outnumber *corpus-driven,* and *corpus-driven* imply that most researchers apply a corpus-based approach. Or it may be assumed that they reject the dichotomy between corpus-based and corpus-driven and call their research corpus-based. For example, McEnery and Hardie (2011) prefer to describe all corpus linguistics as a corpus-based approach. They argue in their book, "for those who accept it, the corpus-based versus corpus-driven dichotomy creates a basic, binary distinction, under which most works of corpus linguistic research can be sorted into one or the other group; however, the researchers' perspective rejects the notion that the corpus itself has a theoretical status, and thus also rejects the binary distinction between corpus-based and corpus-driven linguistics" (p.6).

Swales' CARS model was frequently mentioned, particularly between 2005 and 2010. According to Eak-in (2015), Swales' CARS model is "well-known and has been playing an important role in the analysis of academic texts and been applied widely to explore rhetorical structures of academic work and a part of research articles in various fields" (p.9).

In addition, the corpus research related to RAs seems prevalent in all three periods but tends to be most favorable between 2011 and 2015, including business and academic research. Nesi (2013) notes that "many ESP corpora are made up of research articles, partly because there is a long ESP tradition of research article analysis, dating from Swales' original work on article introductions (1981), and partly because research articles are readily accessible in electronic format and can easily be selected according to learners' specific disciplines and fields" (p.441).

Between 2011 and 2015, lexical *bundles* tended to be the favorite area to be examined, which corresponds with Sa-ngiamwibool (2014) 's findings. Still, the results in the corpus seem to be highly associated with research articles rather than journalistic discourse, as Sa-ngiamwibool suggested.

Corresponding with the trends in ESPj and JEAP, researchers in AESPJ also tend to prefer examining written to spoken texts. This might be because a corpus of written texts is much easier to build. In contrast, a corpus of spoken texts involves a highly labor-intensive process, including several difficulties in data collection (Bowker & Pearson, 2002).

The words related to language and nationality, such as *American, Philippines, Chinese, Vietnamese, Turkish*, and *Persian,* were also revealed between 2005 and 2020. Most words are associated with Asia, Africa, and the Middle East, confirming that most articles published in AESPJ were from these regions. The researchers were intrigued to investigate the use of English and their native language.

The pedagogical application of corpora focuses mainly on the first period between 2005 and 2010 and seems to be under-researched.

Implications of the study

This study demonstrates how corpus linguistics can be applied in trend investigation. It can then provide the applications of such a method for graduate students, researchers, professors*,* and policymakers. For postgraduate students, the corpus-based approach can be selected as one of the methods for identifying a research gap that can gather a large amount of data and reduce reading time. Researchers also employ trend investigation through corpus linguistics to visualize a research gap that needs further investigation. Moreover, in terms of data collection, this approach can be used to guard against criticism of cherry-picking, especially in a qualitative study. For professors, the corpus-based approach can be utilized as a pedagogical tool to promote students' analytical and critical skills since they are required to systematically manipulate a significant amount of data. Most importantly, the authority and policymakers can benefit from the results to gear towards the right direction and push more corpus linguistics research into the ESP domain.

Limitations and recommendations

Since the data of this study were limited to RA abstracts and conclusions, further studies may collect the data from the other parts of the study, such as research methodology and discussion, to obtain multidimensional perspectives. In addition, since this study collected the data from only Asian ESP Journal, further studies can compare trends of the Asian ESP Journal and other ESP journals in other regions. Finally, corpus linguistics for trend analysis should not be only employed for the ESP discipline but also for other fields of research.

References

Anthony, L. (2004). AntConc: A learner and classroom friendly, multi-platform corpus analysis toolkit. *proceedings of IWLeL*, 7-13.

Bortolato, C. (2016). Intertextual Distance of Function Words as a Tool to Detect an Author's Gender: A Corpus-Based Study on Contemporary Italian Literature. *Glottometrics*, *34*, 28-43.

Bowker, L., & Pearson, J. (2002). *Working with specialized language: a practical guide to using corpora*. Routledge.

Charles, M. (2013). English for academic purposes. *The handbook of English for specific purposes*, 137-153.

Coxhead, A. (2013). Vocabulary and ESP. *The handbook of English for specific purposes*, 115-132.

Eak-in, S. (2015). Effects of a corpus-based instructional method on students' learning of abstract writing: A case study of an EAP course for engineering students. *Journal of Studies in the English Language, 10*. 1-41

Fatma, K. A. Y. A., & YAĞIZ, O. (2020). Move analysis of research article abstracts in the field of ELT: A comparative study. *Dil ve Dilbilimi Çalışmaları Dergisi*, *16*(1), 390-404.

Friginal, E., & Hardy, J. (2013). *Corpus-based sociolinguistics: A guide for students*. Routledge.

Gollin-Kies, S. (2014). Methods reported in ESP research articles: A comparative survey of two leading journals. *English for Specific Purposes*, *36*, 27-34.

Handford, M. (2010). What can a corpus tell us about specialist genres. *The Routledge Handbook of Corpus Linguistics. Londres, Nueva York: Routledge*, 255-69.

Hengl, T., & Gould, M. (2002). Rules of thumb for writing research articles. *Enschede.*

Johns, A. M. (2013). The history of English for specific purposes research. *The handbook of English for specific purposes*, *5*, 5-30.

Lesiak-Bielawska, E. D. (2015). Key aspects of ESP materials selection and design. *English for Specific Purposes World*, *46*, 1-26.

McCarten, J. (2010). Corpus-informed course book design. In *The Routledge handbook of corpus linguistics* (pp. 413-427). Routledge.

McCarthy, M., & O'Keeffe, A. (2010). Historical perspective: What are corpora and how have they evolved?. In *The Routledge handbook of corpus linguistics* (pp. 3-13). Routledge.

McEnery, T., & Hardie, A. (2011). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.

Nesi, H. (2013). 21 ESP and Corpus Studies. *The handbook of English for specific purposes*, 407-426.

Ruiying, Y., & Allison, D. (2003). Research articles in applied linguistics: Moving from results to conclusions. *English for specific purposes*, *22*(4), 365-385.

Salmani-Nodoushan, M. A. (2020). English for specific purposes: Traditions, trends, directions. *Studies in English Language and Education*, *7*(1), 247-268.

Sa-ngiamwibool, A. (2014). Recent trends in ASIAN ESP. *Proceedings of the Second International Conference on Education and Language*. Bandar Lampung University.

Scott, M., & Tribble, C. (2006). *Textual patterns: Key words and corpus analysis in language education* (Vol. 22). John Benjamins Publishing.

Tognini Bonelli, E. (2001) Corpus Linguistics at Work. Amsterdam: John Benjamins.

Vo, T. A., & Carter, R. (2010). What can a corpus tell us about creativity?. In *The Routledge handbook of corpus linguistics* (pp. 302-316). Routledge.