References as Cohesive Features of Problem Statement Section in Doctoral Dissertations English Studies in Malaysian University

Nasreen Bhatti

English Department, Faculty of Modern Languages & Communication, Universiti Putra Malaysia, Malaysia
Corresponding Author Email: gs53370@student.upm.edu.my

Ilyana Binti Jalaluddin

English Department, Faculty of Modern Languages & Communication, Universiti Putra Malaysia, Malaysia
Email: ilyana@upm.edu.my

Vahid Nimehchisalem

English Department, Faculty of Modern Languages & Communication, Universiti Putra Malaysia, Malaysia
Email: nimechie@gmail.com

Hui Geng

English Department, College of Foreign Studies, Guilin University of Technology, China
Email: huiyuqiankun@gmail.com

Abstract
One of the most critical sections in a doctoral dissertation is a Problem Statement (henceforth PS). As emphasized this is where most assessors start when trying to comprehend and evaluate the merits of any research study. The PS can then be considered as the basic part of a research study, which may lead to a valuable or, conversely, worthless piece of writing. This type of writing forms a coherent section of the information structured to preserve consistency with the topic and the logical links between the items. A growing body of studies suggests that exploring how certain cohesive devices are employed in a text to create unity provides information that will help L2 students improve cohesion in their writing as they play an essential role in cohesion by organizing and elucidating the relationships of the ideas within the context. One of the cohesive devices that researchers emphasize is references, arguing that ineffective use of referents may weaken a text's cohesion and coherence. This study aims

to investigate references and their different types as cohesive devices and their role in the cohesion of PS writing. For this purpose, the current study used 60 doctoral dissertation Problem Statements published during 2010-2022 from English discipline via Hallidayan Systemic Functional Grammar (Halliday, 1994; Halliday & Matthiessen, 2004). For this purpose, this study employed the discourse analysis method, using a descriptive interpretive approach. TagAnt 2.0.5 software application was used to examine the occurrences and frequency of references in the PSs of doctoral dissertations. Overall, the results suggest that in the present work, the demonstrative references were dominant among the types of references. The definite article '*the*' is greater than the frequency of the other references altogether. Regarding the role of '*the*', the findings indicate that PS writers depend heavily on it to achieve cohesion in their writing. Regarding the Personal references that are used in the corpus of the present study, the findings reported that (*they, theirs*) have been used for cohesion in the PS more than other references in making citations to other studies. Whereas, there was no instance of comparative references found in the corpus of study. In addition, the findings of the present research have implications for PS writers and supervisors to produce cohesive PS. Writers need to be aware of how the references specifically the definite article and the employment of comparative features in their PS writing as cohesive devices contribute to the development of their PS texts.

Keywords: Problem Statement, Doctoral Dissertation, Cohesion, References, Cohesive Device

Introduction
Bearing in mind, that the section of Problem Statement is both descriptive/ informative and argumentative/ persuasive. This type of writing forms a coherent section of the information structured to preserve consistency with the topic and the logical links between the items (Dejica & Superceanu, 2004). When dealing with cognitively demanding texts such as PS, cohesion is even more important. It is argued that writers should use well-structured patterns to ensure the consistency and coherence of their texts. Therefore, the importance of using proper language includes cohesive devices in producing more cohesive and persuasive texts (Ebrahimi, 2014). Student writers may miss cohesive relations if they are unaware of or deficient in this linguistic area, thus having difficulty writing a cohesive PS. Student writers' lack of familiarity with cohesive features and their improper deployment in the text as well often leads them to produce writer-based prose in which the propositional content is not effectively conveyed, thus lowering the overall quality of their PS.

Nevertheless, Chanyoo (2018) argues that among the cohesive devices used in the text, the only device that affects the quality of writing is the use of reference. When students write their texts, they need to employ a number of references as their main cohesive device. In addition, researchers emphasize the references as cohesive devices and argue that an ineffective deployment of referents may erode the cohesion and coherence of a text to achieve cohesion (Crosthwaite, 2017).

Reference mainly refers to the previously mentioned idea at a clausal or sentence level. References are known as a major cohesive device since they are "resources for marking textual status" (Halliday & Matthiessen, 2004, p. 549) that have been previously assigned to the elements of discourse and help the audience in processing the elements through the discourse. In the current study, references refer to the pronouns which function as cohesive elements in the text. Thus, the use of references reflects students' ability to make connections

from the previous segment to the current one by using a word that represents the whole idea, instead of matching word-by-word through the use of reiteration. The claim of the contribution of references to higher writing scores is supported by Alarcon & Morales (2011), who also found that reference had the highest frequency (or 90.67%) of the total grammatical cohesive devices associated with writing quality, followed by conjunction and ellipsis.

It appears from the findings of the various studies on reference usage in writing that many L2 learners utilized references more frequently than other cohesive devices (Afful & Nartey, 2014; Afrianto; 2018; Ampa & Basri, 2019; Dania,2018; Gailea, Syafrizal & Hafipah, 2018; Junina, 2022; Kirana, Mukhrizal & Jayanti, 2020; Puteri, Ermanto & Tahar, 2018; Prasetyaningrum, Asrobi, Surayya & Fikni, 2022; Warna, Wijaya, Sartika & Riznanda, 2019). Research on referential cohesive markers indicates that personal and demonstrative references were used frequently, whereas comparative references were the least used in writing (Cabrejas, 2022; Dania, 2018; Junina, 2022; Ludji, Hambandima & Christiani, 2022; Meisuo, 2000). This may suggest that there is difficulty in using comparative expressions. When it came to the usage of reference ties, student writers generally exhibited three characteristics. First, there were instances where references were used that had neither an explicit referent nor agreement with the text that came right before it. As a result, the reader is left with a lot of uncertainty about the reference. The second category of characteristics pertained to the frequent occurrence of pronoun shifts among students, specifically from singular to plural or first person to second person. The reader may find the reference form system to be inconsistent or even confusing as a result. Suari, Udayana and Parthama (2022) examined the overuse of personal reference. According to Naderi (2014), using personal references like "I" and "we" to express subjective views is discouraged in academic writing, as it is considered neither appropriate nor objective. Instead, personal viewpoints can be conveyed implicitly through other means. On the other hand, Kirana, Mukhrizal and Jayanti (2020) found demonstrative reference is the common type that is used in their research. This finding showed that non-native writers used significantly more demonstrative pronouns than native (Hinkel, 2001, p. 124). The frequency of the definite article, '*the*', significantly increased the rate of occurrence of this category, there were 3802 occurrences of '*the*' making around 74% of referential ties in general, and demonstrative references in particular found in the corpus of this study. The findings of these studies suggest that writers' use of cohesive devices in written texts is influenced by their familiarity with the corresponding cohesive devices in their own language. Cohesive devices that are equivalent in their first language are employed more often, whereas the use of cohesive devices that are not equivalent is limited. However, there are certain limitations to many earlier studies. First, the majority of these studies have concentrated on specific categories of cohesive grammatical devices that students employ most frequently when writing theses. Research places a great deal of emphasis on achieving cohesion, principally through coordinators and subordinating conjunctions. However, less attention is paid to the role of referential pronouns in linking ideas across clausal boundaries (Benell, 2018). Second, these studies largely involve counting frequencies, comparing the overall frequencies, and examining realizations of them in written texts. These studies found that a higher use of references was a good predictor of text organization. They also demonstrate that the type of references used in writing affects its quality, so writing quality cannot be predicted solely by the frequency of their use. However, there has been little discussion about the impact of references on creating discourse. Third, studies examining the use of references within certain sections of thesis writing suggest that

L2 learners tend to use them depending on the type of section. On the other hand, little is known about what a particular set of references is used in the section of Problem Statement, and how they impact the discourse of Problem Statement is unexplored yet.

Methodology
Method
The current study adopted the discourse analysis method, using a descriptive interpretive approach. In this line, the quantitative analysis was applied. The quantitative analysis was used to determine the patterns of occurrences of cohesive features namely references of Problem Statement section in doctoral dissertations. For this purpose, the TagAnt 2.0.5 software was adopted to identify the conjunctions and count their frequency of occurrences. With the help of this software, a wordlist was created and then used to manually identify the occurrences of conjunctions, while any overlaps were removed.

Data Collection
To probe into the references as a cohesive feature of the Problem Statement section, the doctoral dissertations published from 2010 to 2022 formed the corpus of the current research. The doctoral dissertations were selected from the Department of English, Faculty of Modern Languages and Communication to be investigated in this research. A number of inclusion criteria were put in place to build the corpus of doctoral dissertations for the present research. First, all the selected dissertations were written as requirements for the fulfillment of doctoral degrees by ESL students. Second, the selected dissertations had stand-alone Problem Statement sections that were separated from other sections in the introduction chapters because the focus of this research was on the Problem Statement sections, alone. Finally, besides the above-mentioned criteria, publication date and time range had to be decided for selecting PSs. With the view that language is a dynamic and flexible phenomenon (Paltridge & Starfield, 2007), attempts were made to collect up-to-date data. Therefore, the investigated theses were chosen from the time interval of 2010 to 2022. The next step was to analyze the text in light of the adopted analytical framework. Finally, each selected sample was assigned a unique code and a number (e.g. PS1, PS4, and PS10).

Analytical Framework
Although Halliday and Matthiessen (2004) mentioned four elements are involved in the creation of cohesion in discourse (conjunction, ellipsis, reference, and lexical organization), the present study followed Schleppegrell (2004), who suggested conjunctions as cohesion elements of cohesive academic text and more appropriate in textual analysis of academic discourse (Naderi, 2014). Thus, for the cohesive features (i.e., References), the examination of the PS corpus was informed by the framework developed by Halliday and Matthiessen (2004) given below. They categorize reference items into three classes: personal, demonstrative, or definite articles, and comparative, as depicted in Table 1. Biber et al. (1998) note that demonstrative pronouns are common in academic writing as they help establish contextual ties between ideas. Personal references are usually straightforward, referring to a specific person previously mentioned in the text (Halliday & Hassan, 1976). To achieve cohesion, these references must be endophoric that is, related to something mentioned in the text. However, managing the usage of demonstrative pronouns like *this, that, these,* and *those* can pose challenges for L2 academic writers (Crosthwaite, 2017). Failure to effectively

use referents can disrupt the cohesion and coherence of a text if readers struggle to identify the anaphora quickly.

Table 1
Types of Reference Expression

		Nominal group: Head or Pre modifier	Nominal or adverbial group: Sub Modifier	Adverbial group: Head
Co-reference	Personal	Personal pronoun as Thing/Head; possessive determiner as Deictic/Pre modifier or Head	-	-
	Demonstrative	Demonstrative pronoun as Thing/Head; demonstrative determiner as Deictic/Pre modifier or Head	-	Demonstrative adverbs as Head (here, there)
Comparative references	General	Adjective as post-Diectic (same, similar, other, etc.); adjectives such as Epithet	Comparative adverb (identically, similarly, otherwise, etc.) as Sub Modifier in nominal, adverbial group or as pre modifier, Head in adverbial group	
	Specific	Comparative adverb (more, fewer, etc.) as SubModifier of numeral serving as Numerative; comparative adverb (more, less, etc.) as SubModifier of adjectives serving as Epithet (or simply comparative form of that adjectives)	Comparative adverb more, less, etc.) as SubModifier in nominal, adverbial group or as premodifier, in adverbial group (or simply comparative form of adverb)	

(Adapted from: Halliday & Matthiessen, 2004, p. 553)

Data Analysis Tools
The Software tool used in data analysis –TagAnt 2.0.5
To analyze references, TagAnt version 2.0.5 was used to assign parts of speech to each word (and other tokens), such as nouns, references, conjunctions, adjectives, etc., and their frequency of occurrences. TagAnt is a freeware, multi-language tagging tool built on top of the SpaCy natural language processing (NLP) framework. Laurence Anthony (2022) developed this program to generate the following components for each word: WORD, PART-OF-SPEECH (POS) NAME PART-OF-SPEECH (POS) TAG, LEMMA. Users can select language, display component information, choose output format, and adjust line endings using various options.

This computer software was chosen for this research for three reasons. First, it was free software that could be downloaded online and installed on a computer with an excellent user-friendly interface. Second, the software application was practical, providing useful and appropriate tools for the study. In this sense, it helped to make the analysis processes easier and faster as it would have been very difficult and time-consuming to assign parts of speech to cohesive features (conjunctions and references) that could appear in the corpus manually. The whole corpus was analyzed using Tag AntConc 2.0.5 to see occurrences of references as a cohesive feature.

Data Analysis Procedure
In order to answer the research question of the present study, all collected data needed to be analyzed and interpreted. This section explains the methods of data analysis and the instruments used in the current study.

To answer the research question, referential ties were analyzed in doctoral dissertations PSs. To meet the requirement of the software, the PSs were saved in Plain Text format *(*.text file)*. The quantitative analysis of the referential ties was done using the TagAnt 2.0.5 software application. The results were shown as an output display. For example, in the following figure 1 is a screenshot showing the referential ties *them* displayed as PRON.
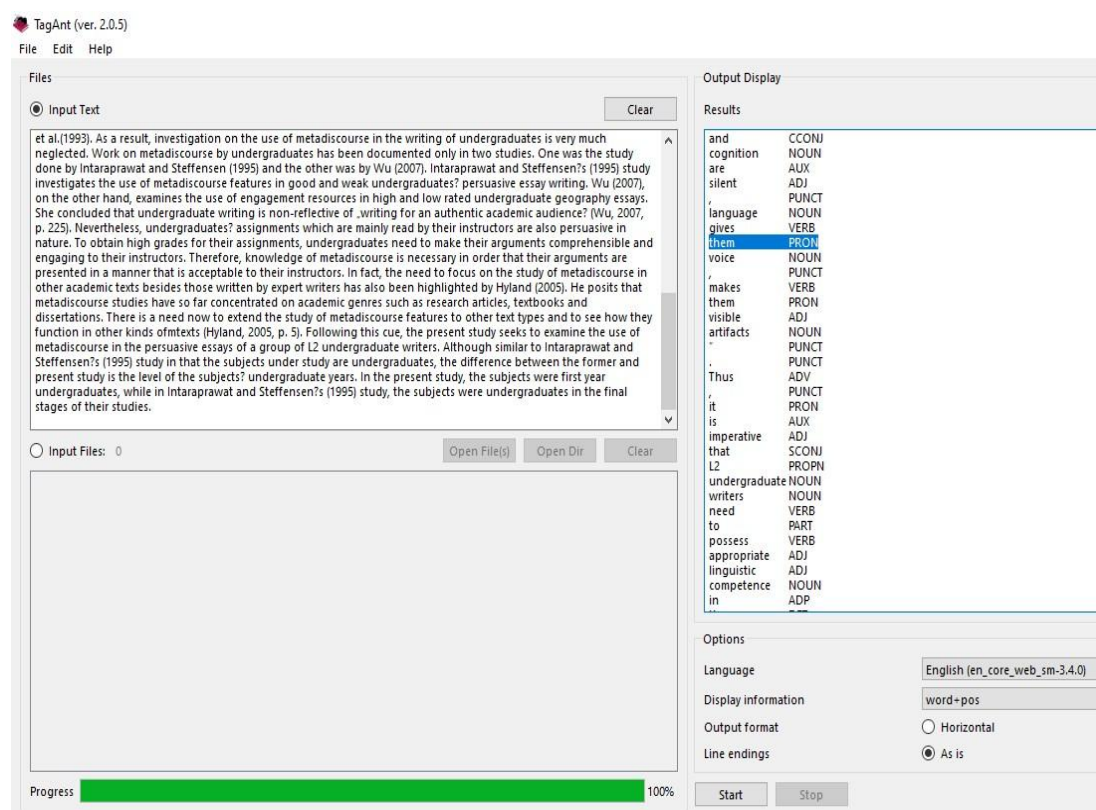

Figure 1: A screenshot showing the identification of references

Figure 2: A screenshot showing the Excel file

The results were saved in an EXCEL file. Figure 2 is a screenshot showing referential ties clearly. As mentioned earlier, software was used to report the occurrences and frequencies of references; however, their functions were analyzed manually. A manual investigation of the referential ties in the PS corpus was performed by counting the referential ties in accordance with the well-developed taxonomy of references developed by Halliday (1994) and Halliday and Matthiessen (2004). Texts of the PS corpus were examined to identify the references. All the instances were analyzed manually using EXCEL (repeating twice as suggested by Naderi, 2014). The total frequency of each class as well as the frequency of each reference was determined. After identification, reference items in each text were classified according to Halliday and Matthiessen (2004) into three classes of references: personal, demonstrative, and comparative. Based on their total number of occurrences, frequencies, and percentages of each referent and each category were determined. The final list was saved in a separate file with a Code name (e.g. *Personal references, Demonstrative references, and Comparative references*) as shown in the figures below.

Pers. REF	he	him	she	her	they	them	it	his	hers	its	their	
2010	6	1		1	3	1		13		8	2	35
2011 (2)					1	1				1		3
2011 (24)								1			1	2
2011 (13)	2	1	3		15	7	1	13		1	29	72
2011 (43)							1				3	4
2012(2)	1							1			6	8
2012 (16)					6	2					9	17
2012 (21)					7						8	15
2012 (22)												
2012 (32)					1	1				3	8	13
2013 (22)	3				8						7	18
2013 (56)	3					2					4	9
2014 (4)					1						2	3
2014 (8)	1				4	1				3	7	16
2014 (15)					9	1		1		2	5	18
2014 (24)					4	3					2	9
2014 (30)					2		1			3	7	13
2014 (64)					11	2				1	11	25
2015 (39)	7	1	3	10			1	8		1	3	34
2015 (46)	1			2	1						4	8
2015 (54)			2		3	1	1	1			5	13

Figure 3: A screenshot showing Personal References

Demonstrative	this	these	that	those	it	the	here	there	W
2010	2		6	2	8	48		9	75
2011 (2)	4	4	4	1	4	18		5	40
2011 (24)	3	5	2		1	58		2	71
2011 (13)	14	5	7	2	8	124		1	161
2011 (43)		1	2		1	25		3	32
2012(2)	2	3	1		2	20			28
2012 (16)	5	4	5		1	53		2	70
2012 (21)	2	2	1		2	18		3	28
2012 (22)									
2012 (32)	7	9	4		2	72		3	97
2013 (22)	11		8	2	8	51		5	85
2013 (56)	8	4	4	2	5	54		2	79
2014 (4)	2	1				28			31
2014 (8)	9	2	2		2	41			56
2014 (15)	21	4	5		5	103		2	140
2014 (24)	6	6	19		1	67		1	100
2014 (30)	6	1	5		4	48		1	65
2014 (64)	7	3	6	2	7	60		8	93
2015 (39)	5	4	7		2	97			115
2015 (46)			3	1		54			58
2015 (54)	2	2	3	1	8	52		4	72

Figure 4: A screenshot showing Demonstrative References

X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH
Comparative	same	similar	other	identically	similarly	otherwise		fewer	more	les
2010	NA									
2011 (2)	NA									
2011 (24)	NA									
2011 (13)	NA									
2011 (43)	NA									
2012(2)	NA									
2012 (16)										
2012 (21)										
2012 (22)										
2012 (32)	NA									
2013 (22)										
2013 (56)										
2014 (4)	NA									
2014 (8)	NA									
2014 (15)										
2014 (24)	NA									
2014 (30)										
2014 (64)										
2015 (39)	NA									
2015 (46)	NA									
2015 (54)										

Figure 5: A screenshot showing Comparative References

Results and Discussion

Referential ties as another cohesive element of the academic discourse were investigated in the corpus of this study. Generally, there were 6318 referential ties found in the corpus of the present study. As it has been mentioned in the previous chapter, Hallidayan systemic functional Grammar (Halliday & Matthiessen, 2004; Halliday, 1994) classifies referential ties into three categories: personal references, Demonstrative references, and Comparative references. The distribution of references was not similar in 60 PSs since many of them limited their usage of the references to one or two categories of references however, there was no instance of comparative reference found in the corpus of study.

Predictably, the most common type of reference was demonstrative. This could be a result of the remarkably high frequency of the definite article '*the*,' which in Hallidayan Grammar is regarded as a demonstrative reference (Halliday & Matthiessen, 2004; Halliday, 1994). The definite article occurred 3802 times in the PSs of doctoral dissertations which is greater than the frequency of the other references altogether (Table 2).

Table 2
Distribution of the definite article

Total references	6318
The	3802
Other references	2516

The frequencies and percentages of the categories of references are presented in Table 3. From the table, it is obvious that the use of demonstrative references was so high that it covered 82% of the total references. Only, 17.9% references of in the text under analysis belonged to personal references. Whereas, there was no instance of comparative references in the corpus of study.

Table 3
Distribution of references in the three categories

References.	Frequency	Percentage
Demonstrative	5154	82%
Personal	1129	17.9%
Comparative	0	0%
Total	6283	100 %

The following Table 4 shows the personal references used in the corpus of the present study. There are 10 types of personal references used in the corpus of the study. Personal references (*they, theirs*) have been used more than other references in making citations to other studies. The current study indicated 431 occurrences of *their* in the corpus as the highest number of personal references making 38% of total personal references. The second highest personal reference is *they* with 351 occurrences in the corpus. There are other instances of references *them, he, she, her, him, her, hers, it, its* in the corpus with lower occurrences in the corpus.

Table 4
Distribution of Personal References

Personal References	Frequency	Percentage
Their	431	38.1%
They	351	31%
Its	74	6.5%
It	72	6.3%
Them	70	6.2%
His	48	4.2%
He	40	3.5%
Her	20	1.7%
She	17	1.5%
Him	6	0.5%
Hers	0	0%
Total	**1129**	**100%**

Consistent with the findings of current literature findings (Akbar, Mujiyanto & Sutopo, 2022; Baybay, 2022; Cabrejas 2022, Khairunisa & Savitri, 2018; Warna et al., 2019) which found '*their*' and '*they*' as the most prominent references, in current study, PS writers display a tendency to use *they* and *their* for cohesion in Problem Statement. However, on the contrary, there are studies that found the personal references *I*, *me*, and *him* as widely used, mostly to indicate the writer's position in the writing **(**He, 2020; Khairunisa & Savitri, 2018). These findings indicate that the writers tend to be more flexible, employing the personal pronoun which may have the danger of resulting in the frequent shifting of pronominals (Zhang, 2000). In contrast, PS writers in this study seem to be more conservative, usually sticking to the use of the third-person plural *they*. It indicates that PS writers may use an objective and impersonal tone in their Problem Statement. Understandably, these writers try to avoid personal references like '*I*' and '*we*' to indicate anything subjective. Naderi (2014) states that authors of Academic writing avoid first-person pronouns ('*I*' and '*we*') since it is believed that it is not honest and suitable to express personal views in an academic context while the idea could be said in other implicit ways which cover any personal point of view. It is also important to note that the use of '*they*' and '*their*' in the PS maintains the reference consistency that avoids causing cohesion to be undermined.

It also demonstrates that the pronouns '*them*,' '*they*', and '*their*' are used to refer to the participants of the research investigation. This overuse of the personal pronouns in reference suggests that instead of repeating a noun monotonously, the writers replace them with pronouns which helps in the cohesion of the Problem Statements. As stated by Halliday, these references are commonly found in research writing (Afful & Nartey, 2014).

Henceforth, it might suggest that the predominant use of reference (*they* and *their*) seemed to be a characteristic of text found in the doctoral dissertation Problem Statement section of the study. In other words, the majority of PS sections shared the same cohesive device to create a written cohesive text.

Table 5

Distribution of Demonstrative References

Demonstrative References	Frequency	Percentage
The	3802	73.7%
This	406	7.8%
That	320	6.2%
It	212	4.1%
These	191	3.7%
There	189	3.6%
Those	33	0.6%
Here	1	0.1%
Total	**5154**	**100%**

Table 5 shows the findings regarding the demonstrative references of this study. Out of four grammatical cohesion types, demonstrative reference, in particular, is the most frequently used device. The use of 'the' as a demonstrative device is more common than any other. It appears in over half of all reference devices used. The high percentage of reference occurrence in this research is dominated by the use of demonstrative reference '*the*' (3802). This finding is in line with several studies (Akber, Mujiyanto & Sutopo, 2022; Alshalan, 2019; Baybay, 2022; Imane, 2020; Kirana, Mukhrizal & Jayanti, 2020; Mindeel, 2021; Mwinyi, 2017; Prasetyaningrum, Asrobi, Surayya & Fikni, 2022). These studies found demonstrative reference as the common type used in this research. This may be explained by the fact that, in contrast to previous research on cohesion, demonstratives constituted the majority of reference devices in this study, closely followed by pronominal reference and limited use of comparative reference. This finding, while contradicting previous research, is consistent with Hinkel's (2001) finding which demonstrated that non-native writers employed considerably more demonstrative pronouns than native writers (p. 124). The frequency of the definite article, '*the*', significantly increased the rate of occurrence of this category. There were 3802 occurrences of '*the*' making around 74% of referential ties in general, and demonstrative references in particular found in the corpus of this study.

The findings show that PS writers significantly depend on '*the*' to achieve cohesion in their writing. This phenomenon is primarily the result of students' overuse of the exophoric demonstrative pronoun 'the' in their writing. According to Tsareva (2010, p.14), demonstratives and definite articles often refer exophorically to something within the context of a situation. "Exophoric reference requires the reader to retrieve the information outside of the writing being read" (Halliday & Hasan, 1976, p.33). In order to illustrate the situation in which excessive exophoric reference takes place, they define the term "language-in-action" (p.34). They contend that when language-in-action occurs, it does not necessarily mean the discourse being produced is ungrammatical, simplified, or incomplete. Furthermore, they point out that language-in-action is usually context-dependent. In addition, as noted by Azzouz (2009), many writers probably do not realize the function of '*the*'. They are aware that '*the*' functions as an article, however, student writers are not aware of its other function as a cohesive device. It could be said that there is insufficient understanding of definiteness and the definite article. Overuse and misuse of the definite article alongside with underuse of the

other referential cohesive words resulted in extensive repetition of the definite article. Therefore, the area of definiteness needs development and improvement, especially in the case of PS writing.

In the related investigations, it appeared that the reference '*it*' was employed personally and mostly in the demonstrative position. Table 6 describes the applications and functions of '*it*' more precisely both as personal and demonstrative references.

Table 6
Distribution of 'It' in the categories

It	Frequency
Demonstrative	212
Personal	72
Total	284

As can be seen in Tables 5 and 6, '*it*' can be used both as a personal and demonstrative reference such as in:

Example: *"Facebook is the most visited website among other social network sites, and **it** has become an integral part of millions of users' different aspects of life."*
And as a demonstrative reference such as in:
Example: "It *is necessary to expand the scope of …."*

In line with the findings of the current literature, in deploying *it* as a cohesive device L2 writers face challenges. It appears that L2 writers use *it* for extended reference, rather than just referring to a single inanimate object, i.e. to a clause or multiple clauses. A key feature of cohesion in English texts is effective extended reference (Halliday & Hassan, 1976). The pronoun '*it*' may be used to refer to a clause, multiple clauses, or even the theme of a paragraph in certain cases where it appears as an unclear reference. Similar to such references are inconsistent with conventional English usage and may disrupt the cohesion of the writing leaving the reader confused. There seems to be a problem with effectively referencing with '*it*' and a consequent loss of cohesiveness which is consistent with the findings of Kim (2012). For the two possible references, '*that*' and '*there*' that could appear in other functions, a qualitative investigation was conducted. As stated by Halliday and Matthiessen (2004) '*there*' is only a demonstrative reference only when it is used as an adjunct, as the opposite of *here,* so the other functions that it may possess cannot be considered as demonstrative references. As determined by qualitative analysis there is no usage of '*there*' in reference positions in the corpus of this study because all the uses of '*there*' were located in an initial position of a sentence or a clause as Existential '*there*'. For example:
Example: "There *is a domino effect that when a university does not have or does not care to have an integrity policy, the teachers and students will also not care about plagiarism."*

It is hard to say what account PSs writers' preference for using '*there*' in an initial position of a sentence or a clause as Existential and not using in reference position, it is a distinct difference. '*that*' can also appear in different functions and positions as a subordinators,

relative pronoun, and complement (Hinkel, 2001). For example, the following sentence (from the texts of the present study) includes two different applications of '*that*'.

Example: *"Direct teacher feedback that focuses heavily on grammar accuracy is not necessarily always effective."*

Example: *"The findings of their study indicate that visual perception is more linked to lust, desire, offence and supervision rather than to knowledge and intellection."*

The first '*that*' shown underlined, is a relative pronoun, while the second '*that*', is a subordinating conjunction. So, a qualitative investigation is needed to find the frequency of '*that*' as a pronoun or conjunction. The results of such examination indicated 546 occurrences of '*that*' employed in subordinating conjunctions and 213 occurrences of '*that*' as relative pronoun positions in the PSs. This may reflect the tendency among PS writers to employ the demonstrative adjective of distant position (that) in object cases and the demonstrative of close proximity (this) in subject positions. These findings for demonstrative referencing (*this that, these* and *those*) also align with Hinkel's (2001) findings which show that these pronouns are used to refer to broader contexts or ideas without clear textual antecedents. This finding may suggest that these terms have multiple functions in English syntax showing greater flexibility in English than they realize, cohesion was generally not impeded.

Regarding comparative references, in contrast to the findings of the current literature (Cabrejas, 2022; Prasetyaningrum, Asrobi, Surayya & Fikni, 2022; Warna et al. 2019), there was no instance of the use of comparatives to make reference to other elements in order to form a cohesive tie in PS corpus of this study. This finding is in line with several studies (Afful & Nartey, 2014; Khairunisa & Savitri, 2018; Mwinyi, 2017; Ong, 2011). This finding indicates the most striking feature was the underuse of comparatives. Among the total number of 6283 referential ties used in the 60 PSs, no comparatives occurred. This could indicate that the writers were not confident when employing comparatives in their writing, or it could be partially the consequence of their ignorance about how to use them as a point of reference in the Problem Statements under study. This finding might also imply that the authors lack adequate knowledge regarding the utilization of how comparative features are used in PS writing or the fact that the Problem Statements under study do not require them. Comparative could have been more effective in explaining the referential ties observed in this study since the texts under investigation were academic writing which have natural connections with relations, and comparisons. Therefore, this area needs more attention (Naderi, 2014).

Overall, based on the analysis of the PSs, among the four types of references, dominating is the demonstrative pronoun *the* used as reference cohesion in the PS. On the other hand, personal pronouns *they,* and *their* used as references are also found in the selected PSs. The current study reveals these three references as the most significant referential features of the Problem Statement. Although higher-quality writing tended to use personal and demonstrative references and temporal conjunctions less and comparative references more, high-quality writing only had significantly greater use of comparative references than low-quality writing. The more proficient writers used demonstrative pronouns significantly more than less proficient students (Chou & Shin, 2014; Liu, 2021).

Conclusion

Overall, the results suggest that predictably, demonstrative references were dominant among the types of references. The use of demonstrative references was so high that it covered 82% of the total references. Regarding the role of '*the*,' the findings indicate that PS writers depend heavily on it to achieve cohesion in their writing. Regarding the Personal references that are used in the corpus of the present study, the findings reported that (*they, theirs*) have been used for cohesion in the Problem Statement more than other references in making citations to other studies. It is also important to note that the use of '*they*' and '*their*' in the PS maintains the reference consistency that avoids causing cohesion to be undermined. Whereas, there was no instance of comparative references found in the corpus of study.

The findings suggest implications for PS writers and supervisors. To produce cohesive Text, PS writers need an awareness of the communicative purpose performed by the Problem Statement in the doctoral dissertations. Additionally, they need to be aware of the linguistic resources that could support them to realize those communicative purposes. For non-native speakers whose writings "are not perceived as appropriate" for professional activity, this is especially crucial (Martínez, 2003, p.104). Therefore, to improve the cohesion of the Problem Statement writing, writers—particularly ESL researchers—should focus especially on the references as cohesive aspects.

The present study contributes by widening the existing body of knowledge of written academic discourse and offering detailed information on the forms and functions of the references as cohesive features in Problem Statements in PhD dissertations based on Hallidayan Systemic Functional Grammar (Halliday 1994, Halliday & Matthiessen, 2013) as its theoretical framework. From the point of view of preparing individuals for their graduate studies here in Malaysian universities, this research has provided insights into the kinds of language that are currently produced as Problem Statements. It will be more important that linguistic features of the existing written discourse of PS can be considered by the teachers of the ESP courses of Malaysian universities and help the tutors to focus on the weak and problematic areas such as overuse of the definite article '*the*'. Since the development of academic writing is fundamental for progress in academic careers. Therefore, a description of the academic writings of postgraduate students can help in improvements of the academic writings of both current and future doctoral students.

References

Afful, J. B. A., & Nartey, M. (2014). Cohesion in the abstracts of undergraduate dissertations: An intra-disciplinary study in a Ghanaian university. *Journal of ELT and Applied Linguistics (JELTAL)*, *2*(1), 93-108.

Afrianto, A. (2017). Grammatical Cohesion in Students' Writing: a Case at universitas Teknokrat Indonesia. *LEKSEMA: Jurnal Bahasa Dan Sastra, 2*(2), 97.

Akbar, K. P., Mujiyanto, J., & Sutopo, D. (2022). Assessing Grammatical Cohesive Devices in the Findings and Discussion of Research Articles by Graduate Students. *English Education Journal, 12*(3), 427-436.

Alshalan, K. F. (2019). *A Systemic Functional Linguistic Analysis of Cohesion and the Writing Quality of Saudi Female EFL Undergraduate Students*. [Master Thesis, Al-Imam Mohammad Ibn Saud Islamic University].

Ampa, A., & Basri, D. (2019). Lexical and grammatical cohesions in the students' essay writing as the English productive skills. *Journal of Physics: Conference Series, 1339*(1), 1-6.

Azzouz, B. (2009). *A discourse analysis of grammatical cohesion in student's writing.* [Unpublished MA Dissertation, University of Algeria].

Baybay, J. T. (2022). Grammatical Cohesion in SHS Students' Research Introductions. *International Multidisciplinary Research Journal, 4*(3), 1 -11.

Benell, T., P. (2018). *Referencing For Cohesion in L2 Academic Writing: A Corpus Analysis.* [Master thesis, İhsan Doğramaci Bilkent University Ankara].

Biber, D. S., Johansson, G., Leech, S., Conrad, & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. Pearson Education.

Cabrejas, M. M. (2022). Cohesive Devices on the Abstracts of Select Research Articles Presented in the International Academic Forum (AIFOR): A Discourse Analysis. *British Journal of Multidisciplinary and Advanced Studies*, *3*(1), 22–41.

Chanyoo, N. (2018). Cohesive Devices and Academic Writing Quality of Thai Undergraduate Students. *Journal of Language Teaching and Research*, *9*(5), 994-1001.

Crosthwaite, P. D. (2017). Managing referential movement in Asian L2 writing: Implications for pedagogy. *Writing & Pedagogy*, *8*(3), 537-558.

Dania, R. (2018). Cohesion in the Abstract of the Theses Written by Undergraduate Students of English Education Program. *Tell-Us Journal*, *4*(2), 141-157.

Dejica, D., & Superceanu, R. (2004). Thematic-Structure Analysis of the Section Statement of Problem in Project Proposals*. Transactions on Modern Languages*, *3* (2004), 45-52.

Ebrahimi, S. F. (2014). *Thematicity in English academic research articles across Disciplines in hard and soft sciences.* [PhD thesis, University Putra Malaysia].

Gailea, N., Syafrizal, S., & Hafipah, A. (2018). *The Analysis of Cohesive Devices in Students' Writing Discussion Text. The Journal of English Literacy Education, 5*(2), 88-98.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*.Longman.

Halliday, M. A. K. (1994). *An introduction to functional grammar* (2nd ed.). Edward Arnold.

Halliday, M. A. K., & *Matthiessen*, C. M. I. M. (*2004*). *An Introduction to Functional Grammar.* (3rd ed.). Routledge.

He, Z. (2020). Cohesion in Academic Writing: A Comparison of Essays in English Written by L1 and L2 University Students. *Theory and Practice in Language Studies*, *10*(7), 761-770.

Hinkel, E. (2001). Matters of cohesion in L2 academic texts. *Applied Language Learning, 12*(2), 111-32.

Imane, A. (2020) *a Corpus-Based Analysis of Grammatical Cohesive Devices in Master Thesis Abstracts*. [Master's Thesis, Oum El Bouaghi University].

Junina, A. K. (2022). Grammatical Cohesion in Argumentative Essays by International EAP Learners in New Zealand. *International Journal of Linguistics, Literature and Translation*, *5*(2), 98–108.

Khairunisa, N. C., & Savitri, W. E. (2018). Analysis on Cohesive Devices Errors in EFL University Students' Argumentative Composition. *Retain*, *6*(1), 70-77.

Kim, Y. (2012). An analysis of referential use in Korean EFL learners' argumentative essays. *English Teaching, 67*(3), 181-204.

Kirana, R. P. Mukhrizal, M., & Jayanti, F. G. (2020). Types of Lexical Cohesion and Grammatical Cohesion in Thesis Abstracts. *Jadila: Journal of Development and Innovation in Language and Literature Education*, *1*(1), 57-68.

Ludji, I., Hambandima, E. S. N., & Christiani, Y. N. (2022). Cohesive devices used in students' argumentative essay; A discourse analysis. *VELES: Voices of English Language Education Society, 6*(1), 14-24.

Meisuo, Z. (2000). Cohesive Features in the Expository Writing of Undergraduates in Two Chinese Universities. *RELC Journal*, *31*(1), 61-95.

Mindeel, E. A. K. (2021). *A study of selected Iraqi theses in terms of cohesion*. [Master's thesis, Karabuk University].

Mwinyi, I. K. (2017). *Effective use of cohesive devices in essay writing by students at the State University of Zanzibar*. [Master's thesis, The University of Dodoma].

Naderi, S. (2014). *A systemic functional approach to academic writing of Iranian graduate students.* [PhD. Dissertation, University Kebangsaan Malaysia].

Ong, J. (2011). Investigating the use of cohesive devices by Chinese EFL learners. *The Asian EFL Journal Quarterly*, *11*(3), 42-65.

Paltridge, B., & Starfield, S. (2007). *Thesis and dissertation writing in second language: A handbook for supervisors.* Routledge.

Prasetyaningrum, A., & Asrobi, M., Surayya, A., & Fikni, Z. (2022). Grammatical cohesion in students' undergraduate thesis. *Premise Journal of English Education*, *11*(2), 297.

Puteri, A., Ermanto, E., & Tahar, H. E. (2018). *Grammatical Cohesion in the Thesis*. Proceeding of International Conference on language, literature, and education (ICLLE) 263, 415-42.

Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective.* Lawrence Erlbaum Associates Publishers.

Suari, N. K. L. A., Udayana, I. N., & Parthama, I. G. N. (2022). The Effects of Cohesive Devices Used in the Background of the Study of English Department Students' Undergraduate Theses. Pustaka: Jurnal Ilmu-Ilmu Budaya**,** *22*(2), 107-113.

Tsareva, A. (2010). *Grammatical cohesion in argumentative essays by Norwegian and Russian learners.* [Unpublished Master's thesis, The University of Oslo].

Warna, D., Wijaya, B., Sartika, E., & Riznanda, W. A. (2019). Cohesive Devices Used in Argumentative Essays of Undergraduate EFL Students in Indonesia. *ENGLISH FRANCA: Academic Journal of English Language and Education*, *3*(2), 125-148.

Zhang, M. (2000). Cohesive features in the expository writing of undergraduates in two Chinese universities. *RELC Journal*, *31*, 61–95.