Linguistic Characteristics In The Research Background And Problem Statement In The Doctoral Thesis Of Arabic EFL Students In Malaysia

Monther Al Jarrah, Kais A Kadhim

[1,2]English Language Centre (PPAL) Universiti Malaysia Terengganu, Malaysia
monther88jarrah@gmail.com, kais.kadhim@umt.edu.my

Abstract

This study explores the linguistic features within the move structures of the research background and problem statement sections in Chapter One of PhD theses written by Arab students in Malaysia. The analysis focuses on four key language elements: verb tense, modality, voice (active and passive), and personal pronouns. A corpus of 30 theses was analyzed to determine common linguistic patterns, revealing significant insights into the language choices of these academic writers. The study employs concordance and hand-tagged analysis to examine verb tense and voice, with manual verification for modality and personal pronouns to ensure accuracy. Eight distinct verb tenses were identified, with the present simple tense being the most frequently used, representing 55.30% of the total instances. This indicates a preference for focusing on current knowledge and ongoing research. In contrast, the past simple and future simple tenses were used to describe past methodologies and future implications, respectively. Modal auxiliaries play a crucial role in conveying sentiments and opinions. The study identified seven modal verbs, with 'can' being the most commonly used, suggesting a focus on possibility over obligation. The analysis of voice patterns revealed a higher occurrence of active voice (79.66%), especially in moves that conveyed up-to-date knowledge. In contrast, passive voice was more common in methodological descriptions, indicating a focus on the actions or subjects rather than the actors. The study also examined the usage of personal pronouns, finding a higher frequency of third-person pronouns, suggesting a more objective narrative style. These findings offer valuable insights into the linguistic choices of Arab EFL students in academic writing, providing a deeper understanding of their communicative intentions and alignment with academic discourse conventions. The results can guide future research on linguistic patterns in academic writing and help educators support EFL students in improving their writing skills.
Keywords: Linguistics; Arab; EFL, Move Structures

INTRODUCTION

The development of research questions is interrogative and reflective in nature, enabling researchers to refine their investigation and come up with new plans. Research questions are developed from self-based inquiries which lead to the formation of subsequent, subordinated, subsidiary, or grounded questions and ultimately the final overarching research questions (Rissanen, 1992;White, 2017; Creswell, 2014; Maxwell, 2013; Sunderland, 2010; Agee, 2009; Andrews, 2003).

According to Agee (2009) and Sunderland (2010), the ever-present reflective process is motivated by self-based inquiries including: "Why would I want to examine

this topic?" "What is the answer I'm looking for?" "What revelations am I expecting to attain?" "What information do I need to find the answers?" Such questions facilitate the formation of the initial focus, leading to the development of sub-questions that narrow down the focus and help navigate the data collection and analysis process. Agee (2009) further explained that as the subject matter becomes clearer, more additional questions would emerge. Such interactive inquiry facilitates the initial focus and leads to the development of key decisions for addressing the research problem.

Maxwell (2013) also asserted the perpetual presence of research questions throughout the research, and their continuous evolution from being provisional questions to becoming focused questions, thus forming an interactive inquiry process. But the author also warned against highly-focused research questions as they may result in a narrowed vision which blurs out key research aspects, important data, projected correlations and phenomena, and possible research gaps. Hence, keeping an open mind is pertinent.

Curiosity often motivates inquiries. Kinmond (2012) stated that a research topic may be triggered by an article, a conversation, discussions and debates, or personal experiences (pp. 24-26). Therefore, research questions are triggered by the researcher's own beliefs, interests, and assumptions. Sandberg and Alvesson (2010) stated that questions are more beneficial than answers because the former motivate reflection and intellectual activity whilst the latter often result in full closure (p.2). Campbell (1982, as cited in Sandberg & Alvesson, 2010) supported the significance of innovative research questions, stating that they may lead to the emergence of novel research problems, and ultimately the solution for age-old debates, integration of multiple approaches, and the dismantling of old wisdom and assumptions (p. 23).

Research into academic writing among English as a Foreign Language (EFL) students has uncovered a range of linguistic patterns and challenges. In examining 90 research articles from three applied linguistics journals, Tseng (2011) identified a common 4-move structure with an optional Background move. The usage of verb tenses in this structure showed that past tense appeared in the third and fourth moves, while present tense was prevalent in the first, second, and fifth moves. Several other studies have focused on corpus-based structural linguistics without incorporating move analysis. Abdullah and Noor (2013), Ang et al. (2011), and Ting et al. (2010) explored how native and non-native speakers utilize lexical verbs, verb-noun collocations, and grammatical features in various corpora. Findings indicate that native learners and Malay learners differ in their usage of past tenses, with Malay learners often overusing past tense and underusing past participles. Additionally, grammatical errors such as preposition mistakes, word-form issues, and subject-verb agreement were identified, highlighting significant areas of improvement for EFL students in Malaysia. This introductory analysis underscores the need for a comprehensive examination of linguistic features in academic writing to inform teaching practices and improve EFL students' academic writing proficiency.

Similarly, Tseng (2011) examined 90 research articles derived from three applied linguistics journals and identified the usage of a 4-move structure rather than a five move-structure as Background move (Move 1) is optional. Together with the move analysis, Tseng (2011) investigated verb tense usage in the identified move-structures. Surprisingly, past tense appeared in the third and fourth moves, whilst present tense appeared in the first, second, and fifth moves.

In addition, there exist certain studies based on corpuses wherein a structural linguistics analysis is carried out without the incorporation of move analysis. Abdullah and Noor (2013), Ang et al. (2011), Ting et al. (2010), and Abdullah and Noor (2013) examined how native and non-native speakers utilize lexical verbs and verb-noun collocations in two corpora: WECMEL (Written English Corpus of Malay ESL Learners) and LOCNESS (Louvain Corpus of Native English Essays). It was discovered that in their argumentative writings, both groups of learners used verb infinitive (VVI), a typical form of lexical verb and verb-noun collocation. Additionally, compared to native learners, Malay learners exhibit an overuse of the past tense (VVD) and an underuse of the past participle (VVN). While this study has demonstrated the important role that lexical verbs and verb-noun collocation play in the works of competent writers (i.e., Malay tertiary students who scored distinctions in English), it only provides the opportunity to examine the linguistic features used by competent writers in the Malaysian setting.

Unlike Abdullah and Noor (2013), Janaki, Chitra and Karen (2013) carried out a corpus-based study using MCSAW and found that Malaysian students face problems in using the past tense auxiliary 'Be' as demonstrated by the seven types of errors identified in their argumentative writing involving 'Shift of Tense', 'Agreement', 'Missing Auxiliary Be', 'Wrong Form of Verb', 'Addition and Misformation', and 'Misordering'.

Ang et al. (2011) investigated verb-noun collocational errors in relation to types and sources in the EMAS sub-corpus of 130 written essays. They found that errors in prepositions are rated as the major problem and that, interestingly, students are in favor of including prepositions instead of avoiding them. Ting et al. (2010) examined the grammatical errors made by university students with low spoken English language proficiency. This study indicated that preposition errors constituted 20.67% of the grammatical errors made by the students. It also identified uncertainties in the correct usage of prepositions as a major problem faced by the students. The researchers went on to indicate that the absence of the plural "-s" ending is an indication of grammatical incompetence, and that the students' frequent mistakes with word-form, subject-verb agreement, and tense demonstrate their poor grammar knowledge.

METHOD

A move analysis, a structure linguistic analysis that identifies a move schema, structural linguistics, a frequency analysis, and an analysis of the most frequently occurring structural linguistics in the corpus for each move are all included in the analysis. A corpus-driven framework for the introduction chapter of social science theses was proposed using semi-structured interviews.

In the current study, the interviews were conducted with subject experts to validate the findings. Convenience sampling was considered appropriate for collecting the necessary data for the semi-structured interviews due to the factors of data accessibility and cost-effectiveness. Results from the study on generalizability and sampling in developmental research led Hultsch et al. (2010) to conclude that the findings from convenient structured samples are compatible with those arising from the random sampling of the same group. According to Bauer and Gaskell (2016), the nature of the issue, the number of relevant milieus, and the available resources all influence the number of people that should be interviewed in qualitative research (p. 43). Nonetheless, some overarching factors must be considered, including the finite amount of possible interpretations or versions of reality and the corpus size. Conversely, it was proposed by

Glaser and Strauss (1999) as well as Lincoln and Guba (2011) that the point of saturation or redundancy would determine how many people to interview. The point of saturation, according to Bogdan and Biklen (1998, p. 62), is the point in data collection at which the information gathered starts to become redundant. According to Lincoln and Guba (2011), three to ten respondents are the minimum number of participants needed to achieve data saturation. Creswell (2015) asserted that there are variations in the sample sizes and demographics among qualitative research projects. The third respondent in this current study was where the researcher hit data saturation. The saturation was identified based on the 30 selected theses.

According to the Discourse-Based Interview (DBI), since the goal of a qualitative study is to acquire a deep analysis of a particular phenomenon rather than to generalize the findings to the population, it is common for these studies to focus on only a small number of instances or persons (Hyland, 2007). According to Creswell (2008) and Dornier (2007), it could be appropriate to interview three to ten people. The number of participants in discourse-based interviews has been extremely low. Petric and Harwood (2013) investigated a female student's writing performance in two assignments through a discourse-based interview, while Harwood and Petric (2012) carried out a discourse-based interview with two participants. The present study chose to interview three participants.

Using a non-random sampling frame (Bordens & Abbott, 2008), the purposive sampling method (Trochim, 2006) was used in this study. This research sought the aid of three experts (n = 3) with experience linked to the creation and usage of the theses with the specific goal of gathering information regarding the application of the identified moves in the genre analysis. The best approach to attain the opinions of people with specialized knowledge in their fields and to support the validity of the genre analysis is through expert sampling (Trochim, 2006). The publication and experience of the participants were the researcher's two primary selection factors in this study. Table 3.2 presents the profiles of the respondents.

Table 1. Respondents' Profile

Respondents	Publication /Articles	Experience / Year
1	15	5
2	25	10
3	52	28

Validation of the genre analysis findings is highly necessary. Therefore, the subject experts were interviewed to gather significantly reliable data and to answer the research question of: To what extent are the moves used by the social science students in their theses? Three subject experts were interviewed. As the subject experts are in charge of writing the academic papers and because it is important to comprehend their cognitive selection of language elements when writing the papers, they were hence selected for the interviews.

The data gathered from the genre analysis and the interviews were analyzed using the combined qualitative-quantitative method. However, this research is generally placed under the perspectives of the qualitative approach. The following sections elaborate on the research instruments and data analysis method so as to address the research questions.

A comparison between Swales' original CARS framework and the framework identified in this study was conducted. An intensive description is provided pertaining to

the differences which is further supported by incorporating the experts' opinions during the interview.

The linguistics features were extracted from the Moves. Next, every pertinent speech fragment was examined within the framework of its lexico-grammatical frames or discourse functions. For instance, Ädel (2006) used a thick method to investigate certain elements of metadiscoursal. The thin approach, on the other hand, is quantitative and makes use of a set of markers that are intrinsically meta discursive. This method works well for corpus research where a lot of data is examined. Hyland (2005) employed a thin approach in his quantitative study. The current study's methodology is quite thin because it relies heavily on earlier research to identify the occurrences of the meta discourse and meta function markers in the dataset. The extraction of the wordlists was carried out using CLAW tagger and analysed using AntConc 3.5.7 version software.

RESULTS AND DISCUSSION

The theses written by the Arab EFL students entailed the usage of tenses, voices, and typology of verbs in the moves. A writer's proficiency with word choice and linguistic elements that are frequently employed in their target discourse community is demonstrated by their proper use of linguistic components, which also demonstrates their ability to write effectively. A total of 30 theses introductions and problem statements were examined to determine their language characteristics to answer the researcher question, What are the linguistic features used within the move structures of the research background and problem statement sections in Chapter One of the PhD theses written by Arab EFL students in Malaysia?. This study focused on four target language features: voice, modality, verb tense, and personal pronoun. Both concordance and hand-tagged analysis were used in this study.

The active and passive voices as well as verb tenses were examined manually. Conversely, the primary method of identifying modalities and personal pronouns was manual checking to guarantee the accuracy of the results. But the variations in introduction/problem statement types were not taken into account while analyzing these target linguistic traits. The analysis results pertaining to this research question are arranged according to the linguistic aspects, ranging from verb tense, modal verbs, formation of the active and passive voices, to personal pronouns.

Verb Tense

To ascertain their tenses, the finite verbs of each move in the corpus of 30 introductions and problem statements from the PhD theses were examined. The percentage and frequency of the verb tenses present in the moves in Phase I of the corpus are displayed in order of decreasing frequency.

Table 2. Frequency of verb tenses in Phase I

Tense	M 1 (N=350)				M 2 (N=257)			M 3 (N=164)	M 4 (N=135)	M 5 (N=70)		STP (N=145)	Total	%
	S 1A	S 1B	S 1C	S2	S 1A	S 1B	S2			S1	S2			
Present Simple	199	20	5	32	138	17	22	50	40	18	27	52	620	55.30
Past	19	5	1	2	27	6	4	105	83	3	6	13	274	24.4

Simple														4
Future Simple	6	-	-	2	34	1	-	3	1	3	11	73	134	11.95
Present Perfect	28	5	-	8	2	-	1	3	8	1	1	5	62	5.53
Present Continuous	7	2	-	9	4	-	-	1	2	-	-	2	27	2.40
Past Continuous	-	-	-	-	-	-	-	2	-	-	-	-	2	0.17
Present Perfect Continuous	-	-	-	-	-	1	-	-	-	-	-	-	1	0.08
Past Perfect	-	-	-	-	-	-	-	-	1	-	-	-	1	0.08
Total	259	32	6	53	205	25	27	164	135	25	45	145	1,121	100

Note: * M= Move
**S= Sub-move
***N= the total number of occurrences
****STP = Structuring the presentation

Eight distinct verb tenses were employed in Phase I, totaling 1,121 occurrences, as table 2 illustrates. This entailed three past tenses (Past Simple, Past Continuous, and Past Perfect), one future tense (Future Simple), and four present tenses (Present Simple, Present Perfect, Present Continuous, and Present Perfect Continuous). The present tense was the most frequently occurring tense across the board. This indicates that the writers mostly favored using the present tense to achieve their intended communication purpose. Upon examining the verb tenses distribution in depth, the Present Simple was found to be the most often used tense, appearing in 55.30% of the 620 instances. At 32.09% (199 occurrences), it was common in Move 1 Sub-move 1A: Stating current knowledge. Move 1 Sub-move 1C: Extending previous research, however, hardly employed the Present Simple tense. The Past Simple tense, which occurs 24.44% of the time (274 instances), was the second most commonly used verb tense. In Move 3: Describing the methodology, the Past Simple tense was primarily used (38.32%, 105 occurrences) to explain the research methodology. In Move 1 Sub-move 1C: Extending previous research, this tense was the least used. Nonetheless, the fact that the corpus contained very few examples of this sub-move may account for the low frequency of the Past Simple tense in Move 1 Sub-move 1C: Extending previous research. The third most commonly used tense, at 11.95% (134 times), is future simple tense. In the STP move, there were 73 instances of the chosen tense, which had a frequency rate of 54.47%. The tenses that were analysed

Vol. 7 No. 3 / October 2024

the least in the corpus were the Past Perfect and Present Perfect Continuous. These two tenses only happened once each, representing 0.08%.

Modality

One of the most crucial parts of the English language are modal auxiliaries. They can transmit sentiments, opinions on statements, and evaluations of the speaker's penchant (Halliday, 1970; Palmer, 1979; Quirk et al., 1985). The modal auxiliaries in a corpus of 30 introductions / problem statements taken from Arab PhD theses were examined. Table 3 presents the percentage and frequency of occurrence for the modal auxiliaries in Phase I, in descending order.

Table 3. Frequency of Modality in Phase I

Modality	Number of occurrences	Percentage of occurrence
Can	59	55.66
Should	12	11.32
Could	11	10.37
Would	10	9.43
May	7	6.60
Must	4	3.77
Might	3	2.83
Total	**106**	**100**

The phase I corpus entailed seven types of modal verbs: 'can', 'should', 'could', 'would', 'may', 'must', and 'might'. They appeared 106 times overall. But when compared to the corpus' overall number of words, these modal verbs only appeared 0.72% of the time. In terms of verb forms, the corpus had three present forms and four past forms. But there was a significantly higher occurrence of the present modal verbs (66.03%) compared to the past modal verbs (33.96%). The most used modal verb was 'can' (55.66%, 59 instances). The modal verb 'can' was majorly analysed to signal possibilities (73%). Meanwhile, 'should' and 'could' were the second (11.32% or 12 occurrences) and third (10.37% or 11 occurrences) most frequently occurring modal verbs. The least used modal verb was 'might' (2.83% or 3 occurrences). The data suggests that the students tend to evade showing obligations in their writing, which is why the modal verb 'must' had a low rate of instances (3.77% or 4 instances). Actual usage of these modal verbs is exemplified below, presented alphabetically.

Examples:
1. *Can*
   There are a number of ways in which learning technology can be used to enhance English language teaching and learning.
2. *Could*
   Previous studies reported learners believed that translation could help enhance their foreign language skills including vocabulary knowledge.
3. *May*
   Striving to be correct, their written language skills may overshadow their oral capabilities.The 1995 survey originally aimed to identify the reason why the style of teaching and learning in Japan may not sufficiently help students to communicate in real second language situations.
4. *Might*
   The research aims to investigate whether threaded discussion can represent an

important forum for opening up new learning possibilities that might not be achievable in face -to-face classroom alone.

5. *Must*

   If it is used, it must be in literary works and quotations.

   Instead of avoiding the technologies or completely rewriting curriculum and materials for them, teachers must either learn how to effectively integrate traditional materials into a modern setting or investigate how to integrate the new technology into a traditional setting.

6. *Would*

   Seeking to discover how students would respond to a negotiated syllabus, and what impacts a negotiated syllabus can have on learner autonomy, this study followed a group of young adults from hill tribe regions of Burma and Thailand in an entry-level English course.

English in Oman is seen as an important foreign language. It is analysed in both governmental and private organizations. However, it is used more in private associations, as they are linked more closely to international organizations, such as UNESCO, UNICEF and others areas such as oil and gas companies. From this international perspective, people in Omani society see the need to focus more on how English is taught in classrooms. They also see the importance of teaching English as a means for communication and of developing children's abilities to use it authentically - for example, filling in English forms and writing e-mails. Parents see the importance of this target language for their children as 3 it operates as a tool or as a gateway to the future. They spend their time, effort and money to help their children to learn English quickly and effectively. They sometimes send their children to private institutes during holidays for English and computing courses. From this standpoint, English is expected to be taught to reach the expectations of society and parents. These expectations lead to focus on the use of the target language.

They mainly focus, as the researcher sees it, on language production skills, as parents are keen to see their children write and talk well. The use of English in Oman could have been developed into "institutionalized domains", such as business, the media and education (Al-Busaidi, 1995). To prepare future participants to function in these domains, English is taught in government schools from Grade One, while it is taught from Kindergarten One in private schools. English has also become the medium of teaching and training in all private and public higher education/post-secondary institutions throughout the Sultanate (Al-Issa, 2005). It is clear that English is seen as an effective tool for any educational innovation. As such, efforts to improve English Language instruction receive political and economic support from the government, which determines its place on the social "hierarchy" (Al-Issa, 2002). English is considered as a resource for "national development" (Wiley, 1996). "English is also considered as a fundamental tool that facilitates 'Omanisation'" - a gradual and systematic process through which the expatriate labour force is replaced by a qualified Omani one (Al-Issa, 2002). Mastery of the English language is a prerequisite for finding a white-collar job (Al-Busaidi, 1995; Al-Issa, 2002). English is, hence, should be a central to Oman's "continued development" (Nunan, Tyacke, & Walton, 1987, p. 2) and is "a resource for national development as the means for wider communication within the international 4 community" (2). This stresses the need for English language and also a need to specify the objectives of learning it in Oman and other related elements and of course the need to

make these elements easy to be known by different organisations in a form of a framework which this study intents to develop."

Active Voice and Passive Voice

The active and passive voices in the samples were examined. Voice patterns assist in displaying a writer's concentration in addition to physically articulating the meanings of words combined. The active voice pattern is used by writers to highlight objects or people who are acting. Conversely, the passive voice pattern is purposefully employed to bring attention to certain items or behaviors. Table 4.8 presents the identified voice constructs' frequency and percentage in each move in Phase I of the corpus.

Table 4. Identified Voice Constructs In Each Move

Voice	M 1 (N=350)				M 2 (N=257)			M 3 (N= 164)	M 4 (N= 135)	M 5 (N=70)		STP (N= 145)	Tot al	Perc entag e
	S 1A	S 1B	S 1C	S2	S 1A	S 1B	S2			S1	S2			
Active	217	25	5	45	180	23	24	96	109	23	29	117	893	79.66
Passiv e	42	7	1	8	25	2	3	68	26	2	16	28	228	20.33
Total	**259**	**32**	**6**	**53**	**205**	**25**	**27**	**164**	**135**	**25**	**45**	**145**	**1,1 21**	**100**

Note: * M= Move
**S= Sub-move
***N= The total number of occurrences
****STP = Structuring the presentation

Table 4 indicates a total of 1,121 instances of voice patterns, whereby the occurrence of active voice (79.66%, 893 instances) clearly outnumbers that of the passive voice (20.33%, 228 instances) But both the active and passive voice occurred in all the moves and sub-moves in the corpus. A majority of the active voice occurred in Move 1 Sub-move 1A: Stating current knowledge (24.30%, 217 instances), and Move 2 Sub-move 1A: Indicating main features (21.45%, 180 instances). Move 1 saw a high occurrence of active voice due to its function of conveying up-to-date knowledge of the research topic. In contrast, a majority of the passive voice occurred in Move 3: Describing the methodology (20.15%, 68 occurrences) due to its function of describing the actions, people, or things being studied (e.g., variables, research participants, research tools). The following are examples of active voice usage in the sub-move of Stating current knowledge as well as the passive voice in the sub-move of Describing the methodology.

Examples: *Active voice in Move 1 Sub-move 1A: Stating current knowledge*
One factor underlying the low listening and speaking proficiency of learners of English as a foreign language is the discontinuous use of the language. Intensive teaching of these two skills, with emphasis on varieties of activities, is an alternative method to bring about more fruitful outcomes.

Accordingly, an urgent need exists to study Palestinian control practices under the unique politico-economic uncertainty that has dominated the environment for decades, which, in fact, is one of the longest lasting politico-economic uncertainty in the world. In addition, it is also fundamental to understand the role of the Palestinian national culture in influencing the behaviour of top management in the design and use of MCS, especially under such long-term uncertainty that influences the national culture. Therefore, it is fundamentally necessary to examine how Palestinian organizations design and use their

control systems under the influence of their unique environment and culture that has spent decades under the pressure of the politicoeconomic uncertainty. This examination is an essential step toward driving the performance of the Palestinian organizations. In this context, a search of the previous literature reveals that under conditions of environmental uncertainty, effective MCS design can assist in enhancing organizational performance. However, such enhancement is based on the condition that the design harmonizes with the intrinsic and the extrinsic contextual factor to effectively confront the environmental uncertainty surrounding these companies. (Chenhall & Morris, 1986; Govindarajan, 1984; Khandwalla, 1972; Otley, 2012; Simons, 1990, 2013). In fact, many studies have attempted to approach the difficulties of MCS design with respect to its contingent factors (Chenhall, 2003; Chow et al., 1999; Fisher, 1995; Otley & Wilkinson, 1988).

International journals or even some national journals request English manuscripts. Several studies show that non-native English speaking (NNES) writers face difficulties to publish their work in English. Passive voice in Move 3: Describing the methodology From this group, 20 students were randomly selected for an interview about the quality of their experience. They were interviewed in small group of 5 to provide the opportunity to hear the student's voices.

The study: Korean and Chinese students in an English Vocabulary course at a Korean university were divided into two classes. The course's focus was on learning a large quantity of vocabulary. Group A were taught 3 VLS before mid- term exams; group B were taught the same 3 VLS in the second half of the semester. Both groups' vocabulary levels were tested 5 times during the semester, and they answered 3 surveys about vocabulary learning and VLS.

Although GFMIS is perceived as an effective service-delivery tool for the government, its implementation in Jordan faced several operational challenges. First, Jordanian ministries and departments do not have strong leadership that can initiate and put in place a well-functioning system, such as GFMIS (Biggs, 2012; USAID, 2014a). This must have been occasioned by the lack of support from relevant agencies, which are responsible for managing the processes. Second, the increased number of government offices that are using GFMIS against the original estimated number when it was launched (USAID, 2014a; USAID, 2013). This caused system overburden during its implementation.

Personal Pronoun

In order to avoid repeating a noun that has been mentioned twice, a personal pronoun is analysed in its place. Table 4.9 presents the frequency and percentage of personal pronouns in Phase 1 of the corpus. The results are arranged according to the type of pronoun.

Table 5. Personal pronoun usage frequency in Phase I

Personal pronoun		Number of occurrences	Percentage of occurrence
1st person	*I*	10	6.02
	We	26	15.66
2nd person	*You*	20	12.04
3rd person	*He*	4	2.40
	She	1	0.60
	It	61	36.74
	They	44	26.50
Total		**166**	**100**

Table 5 demonstrates that all the types of personal pronouns (first person, second person, and third person) were analysed in Phase 1 of the corpus, occurring 166 times. In terms of frequency, the third-person pronouns showed the highest rate of usage with 110 instances, compared to the first-person pronoun at 20 occurrences and second-person pronoun at 36 instances. 'It' was the most frequently used personal pronoun (36.74%, 61 instances), trailed by 'they' (26.50%, 44 instances). The least used pronoun was 'she' (0.60%, 1 instance).

There was interesting usage of the personal pronouns. First of all, the first-person pronouns were mostly found in the move of Structuring the presentation. 'We' had a higher usage than 'I', i.e., 26 times (72.22%) compared to 10 times (27.77%), respectively. Next, the personal pronoun 'we' was used inclusively and exclusively. Inclusively, 'we' denotes the writer/speaker and addressee, whilst exclusively, it denotes only the writer/speaker and not the readers (Martin, 2003b; Lores, 2006). The inclusive 'we' showed more frequent usage (17 instances, 65.38%) compared to the exclusive form (9 instances, 34.61%). Third, the referential pronoun 'it' occurred 38 times (62.29%) compared to its non-referential form which only occurred 23 times (37.70%). The non-referential 'it' largely appeared in the passive form. Finally, out of the four occurrences of the pronoun 'he', two denotes the researcher". The contextual usage of personal pronouns is exemplified below, in the sequence of first-person pronoun, second-person pronoun, and third-person pronoun:

Fifth, a survey by FRP II evaluation team on GFMIS users has demonstrated that the system is not user-friendly and even slower than the legacy system (USAID, 2014a). Some respondents complained that the system is slow, as it has many windows which give rise to work duplication (USAID, 2014a), and that the system does not provide accurate outputs and results. There is also work delay due to frequent breakdowns which cause problems in service delivery. Sixth, in examining the adoption status of GFMIS in Jordan, Shannak (2015) also identified the issues that hindered its proper implementation, including the end users' opposition to the new system, their poor commitment to attending the necessary training courses, and their infrequent usage of the system. In studying employee acceptance of GFMIS, Sawalha and Abu-Shanab (2015) indicated that GFMIS is not well-developed enough to manage public financial transactions. Meanwhile, reports by the World Bank suggest that the GFMIS has yet to demonstrate substantial benefits for MDA employees (World Bank, 2015; World Bank, 2016a).

First-Person Pronouns

Overall, 8 tenses were found in Phase 1 of the corpus. They can be grouped as present tense (*Present Simple, Present Perfect, Present Continuous*, and *Present Perfect Continuous*), past tense (*Past Simple, Past Continuous*, and *Past Perfect*), and future tense (*Future Simple*). *Present Simple* denotes the most frequently used tense in the corpus, whilst the *Present Perfect Continuous* and *Past Perfect* were the least used. In the corpus, seven modal verbs were identified: *'can', 'should', 'could', 'would', 'may', 'must'*, and *'might'*, with *'can'* being the most used and *'might'* the least used. In terms of voice construction, the passive voice occurred more than the active voice. The active voice was more prominent in *Move 1: Situating the research*, whilst the passive voice in *Move 3: Describing the methodology*. In terms of personal pronouns, the corpus includes the usage of the first person, second person, and third person. Compared to the first-person and

second-person pronouns, the third person was more popular. *'They'* denotes the most prominent personal pronoun whilst *'she'* was the least popular. The referential *'it'* was used more than the non- referential form, which mostly occurred in the passive voice.

This study aims to examine the compatibility between the theses' structural components and Swales' (1990) CARS model and Bhatia's (1993) theses move structure. The findings from Promsin' (2006) study suggested that the moves in engineering theses differ in both settings. It was found that the 20 theses written by Thai students from the School of Engineering and Technology of the Asian Institute of Technology (an international university) displayed the use of all three moves in Swales' CARS model. Nevertheless, Promsin (2006) found some differences between the move structures identified by Swales (1990) and those in the selected theses. For example, Swales' Step 1 in Move 1: Claiming Centrality and Move 2: Establishing a Niche appear to be an optional move used by the Thai students.

On the other hand, the moves identified in the 20 theses written by Thai students from the Faculty of Engineering, Chulalongkorn University (a Thai university) were found to be compatible with Bhatia's (1993) move structure. However, Bhatia's (1993) Move 4: Presenting Conclusion was often omitted. According to Promsin (2006), the students prefer to start their writing using phrases such as "the objective of this study is…" and "This thesis proposes…" which reflect the use of Bhatia's (1993) Move 1: Introducing Purpose. Interestingly, Promsin's (2006) study highlighted the absence of Bhatia's (1993) Move 4: Conclusion which may be a result of the authors' non-adherence to the IMRD (Introduction, Method, Result and Conclusion) move structure as proposed by Bhatia's (1993) traditional four-move theses structure. Quoting Martin (in Promsin, 2006:54) about the unclear division between the Results and Conclusion units from the students' perspective, Promsin (2006) felt that this was because the students were unable to distinguish between the moves for the Results and Conclusion sections.

Within the EAP areas, move analysis and structural linguistics analysis complement one another by concentrating on the general rhetorical patterns of language elements in a range of professional and academic genres. The following section elaborates on corpus-based studies to examine the structural linguistic patterns of academic and professional genres.

CONCLUSION

The examination of the linguistic features within the move structures of the research background and problem statement sections in Chapter One of PhD theses written by Arab EFL students in Malaysia sheds light on the language choices made by these academic writers. Across the 30 theses analyzed, the data reveals distinct patterns in verb tenses, modal verbs, active and passive voice, and personal pronouns. The predominance of the present simple tense, with 55.30% usage, indicates a tendency for these writers to focus on current knowledge and ongoing research in their field. In contrast, the past simple and future simple tenses, which occur with less frequency, are used mainly to describe past methodologies and future implications, respectively. Additionally, modal verbs such as 'can', 'should', and 'could' were identified, suggesting a preference for expressing possibility and conditionality over obligation and necessity.

The analysis of active and passive voice constructs demonstrates that the writers frequently used the active voice to assert statements and communicate actions, while the passive voice was primarily used in methodological descriptions, indicating a shift in

focus to the actions or subjects rather than the doers. The study also found an interesting pattern in the use of personal pronouns, with the third person being the most common, highlighting a more objective and detached narrative style. The results from this study contribute to a broader understanding of the linguistic choices made by Arab EFL students in academic writing and offer insights into how these linguistic patterns reflect the writers' communicative intentions and alignment with established academic discourse structures.

REFERENCE

Abdullah, S., & Noor, N.M. (2013). Contrastive Analysis Of The Use Of Lexical Verbs And Verb-Noun Collocations In Two Learner Corpora: WECMEL vs. LOCNESS. In S. Ishikawa (Ed.), Learner Corpus Studies In Asia And The World (Vol. 1, pp.139-160). Kobe University

Ädel, Annelie. 2006. Metadiscourse in L1 and L2 English. Amsterdam/Philadelphia: John Benjamins

Agee, J. (2009). Developing Qualitative Research Question: A Reflective Process. International Journal of Qualitative Studies in Education, 22(4), 431–447. Bloomsbury Academic & Professional

Al-Busaidi, K. A. (1995). *English in The Labour Market In Multilingual Oman With Special Reference To Omani Employees* (Doctoral dissertation, University of Exeter).

Al-Issa, A. (2002). An Ideological And Discourse Analysis Of English Language Teaching In The Sultanate of Oman (Unpublished doctoral dissertation). University of Queensland, Australia.

Ang, L. H., Hajar, A. R., Tan, K. H., & Khazriyati, S. (2011). Collocations in Malaysian English learners' writing:A corpus-based EA. 3L: The Southeast Asian Journal of English Language Studies, 17 (Special Issue),31-44. http://journalarticle.ukm.my/3043/

Bogdan, R., & Biklen, S. K. (1998). Qualitative Research for Education: An introduction to theories and methods. Boston: Allyn and Bacon, Inc.

Bordens, K.S, and Abbott, B.B. (2008) Research design and methods: a process of approach. (7th Edition). New York: Mc Graw Hill.

Campbell, K. W. (1982). Bayesian Analysis Of Extreme Earthquake Occurrences. Part I. Probabilistic hazard model. *Bulletin of the Seismological Society of America*, *72*(5), 1689-1705.

Chenhall, R. H., & Morris, D. (1986). The Impact Of Structure, Environment, And Interdependence On The Perceived Usefulness Of Management Accounting Systems. The Accounting Review, 61(1), 16-35

Creswell, J. (2015). Research Design: Qualitative, Quantitative and Mixed Methods Approaches. Pearson Education Inc.

Creswell, J.W. (2008). Educational Research: Planning, conducting and evaluating quantitative and qualitative research. (3rd ed.). USA, Boston: Pearson Education.

Glaser, B. G., & Strauss, A. L. (1999). *The Discovery Of Grounded Theory: Strategies For Qualitative Research*. New Brunswick: Aldine Transaction.

Govindarajan, V. (1984). Appropriateness Of Accounting Data In Performance Evaluation: An Empirical Examination Of Environmental Uncertainty As An Intervening Variable. Accounting, Organizations and Society, 9(2), 125-135.

Halliday, M.A.K. (1970). "Modality and Modulation in English" in G.R. Kress (ed). Halliday: System and Function in Language. (76-112). Oxford , OUP

Harwood, N., & Petrić, B. (2012). Performance in The Citing Behavior Of Two Student Writers. *Written Communication*, *29*(1), 55-103.

Hyland, Ken, 2007. Applying A Gloss: Exemplifying And Reformulating In Academic Discourse. Applied Linguistics 28, 266--285.

Hyland, K. (2005). Metadiscourse: Exploring interaction in writing. New York, NY: Continuum.

Khandwalla, P.N. (1972). The Effect Of Different Types Of Competition On The Use Of Management Controls. Journal of Accounting Research, 10, 275-285.

Kinmond, K. 2012. Coming Up With A Research Question. Pp. 23-36 in C. Sullivan, S. Gibson and S. Riley (Eds.), Doing Your Qualitative Psychology Project. London: Sage

Lincoln, Y.S., Lynham, S.A. & Guba, E.G. (2011). Paradigmatic Controversies, Contra-Dictions, And Emerging Confluences, Revisited. In N.K. Denzin & Y.S. Lincoln (eds),The Sage Handbook of Qualitative Resear ch (4th edn; pp. 97–128). ThousandOaks, CA: Sage.

Maxwell,J.A(2013). Qualitive Researches Design An Interactive Approach (uc.ed). Thousand oaks, CA: Sage

Nunan, D., Tyacke, M., & Walton, D. (1987). Philosophy and Guidelines For The Omani English Language School Curriculum. *Muscat, Sultanate of Oman: Ministry of Education and Youth*.

Palmer, (1979). Modality and the English Modals. London: Longman.

Petrić, B., & Harwood, N. (2013). Task requirements, Task Representation, And Self-Reported Citation Functions : An exploratory study of a successful L2 student's writing. Journal of English for Academic Purposes. 12, 110–124

Quirk, R. (1985). A Comprehensive Grammar of English Language. London: Longman

Reeve, J., & Tseng, C. M. (2011). Agency As A Fourth Aspect Of Students' Engagement During Learning Activities. *Contemporary Educational Psychology*, *36*(4), 257-267.

Rissanen, Matti. (1992). *History of Englishes: New Methods And Interpretations In Historical Linguistics*. Mouton de Gruyter.

Simons, R. (2013). Performance measurement and control systems for implementing strategy: Text and Cases: Pearson New International Edition. Upper Saddle River, New Jersey: Pearson Higher

Sunderland, J. (2010). Research question in linguistics. In Litosseliti, L. (ed.), Research methods in linguistics (9–28). London: Continuum International Publishing Group.

Ting, S. H., Mahadhir, M. & Siew, L. C. (2010). Grammatical Errors In Spoken English Of University Students In Oral Communication Course. GEMA Online Journal of Language Studies, 10(11), 53-70

Trochim, W.M. (2006). The Research Methods Knowledge Base. From: http://www. socialresearchmethods.net/kb/.