

# INTRODUCTION

## 1.1. Background of the Study

This study is concerned with creating a corpus containing English extracted from textbooks in the discipline of Islamic religious studies which are used as core references in Indonesian Islamic universities and colleges. The corpus is named the Corpus of Islamic Religious Studies Textbooks (CIRElisST) in Indonesian State Institutes for Islamic Studies (ISIIS). The CIRELISST-ISIIS is then used as the source for creating a technical vocabulary list which is called the Islamic Religious Studies Textbooks Vocabulary (IRSTV) representing features of Islamic values.

The present study sets itself against the backdrop of issues surrounding EFL materials developed locally to meet the needs of the learners of EFL in Islam-related universities and colleges. Under the overarching theory of corpus linguistics and ELT materials development, this research is linked to the shared concept of these two previously mentioned theories. This shared concept is related to the principle that regulates the language to be presented in textbooks/course books. The principles regarding the language content as the key element of any language teaching textbooks and course books on which this study is based, are the importance of authentic and natural language/texts (Tomlinson, 1998, 2008, 2012) . Another principle is the principle that materials should cover most frequent or sufficient samples of language items (Tomlinson, 2010) to maximize learning (Nation, 1993 as cited by Maley 2016; Nation, 2016) which requires sufficient analysis of corpus of L2 (Harwood, 2005; Molavi, Koosha, & Hosseini, 2014). In the same vein, one of the corpus-linguistics' main tenets is that corpus-based studies result in empirical quantitative findings on language features that exist in real or natural setting, i.e. language as used by language communities (McEnery & Hardie, 2012; McEnery & Wilson, 2003). The shared keyword between the principles of materials development and corpus linguistics is 'real or natural language', which is then used to envelope the focus of the present research of creating a specific target corpus and technical word list to be applied in a corpus-informed materials development design for Islam-based tertiary education institutes in Indonesia.

The ELT context of this present study is academic English for undergraduate (freshmen) level students of an Islamic government-owned college in eastern Indonesia. One of the purposes of learning EFL in most Indonesian Islamic universities and colleges managed by the Ministry of Religious Affairs (MORA) is to master English whilst learning and promoting the teachings of Islamic values. This purpose is enhanced by the need for having an authoritative English version of religious (Islamic texts i.e. the Holy Qur'ān and Ḥadīth) that has increased with the growing number of native or non-native English-speaking Muslims (Hassan, 2016). Muslim learners are, therefore, required to be able to use English as one of the tools to communicate Islamic agendas to fellow Muslims as well as to the wider non-Muslim citizens of the world (Abudukeremu, 2010; Rohmah, 2012; Salehi, Davari, & Yunus, 2015). The closest implementation of this purpose in an academic context is the growing needs to produce academic writing texts in English to disseminate the knowledge of Islam (Abudukeremu, 2010). At the same time, the need of vocabulary in reading or listening no matter what level it is is still undebatable classic issue for ESL/EFL learners worldwide, including the ones in the same context as the students of the targeted context of EFL.

To achieve this purpose of EFL learning in this specific context, materials used in Islamic-related (higher) education were designed and produced to present and convey Islamic values and traditions through extensive use of subject-specific texts of Islam-related topics and the use of Islam-related vocabulary (henceforward abbreviated as IRV). Some EFL textbooks compiled by Arsyad (2001) and Kardimin (2011) exemplify the effort to introduce IRV by integrating IRV of several categories (see also Hassan, 2016 for the classification of Islamic-religious Terminologies/IRT).

In the context of EFL materials for students of Indonesian Islamic universities and colleges, there has been a considerable on what kind of vocabulary choice is best to represent the values of Islam in EFL textbooks. The available choices of the versions of the Islam-related vocabulary in English are Arabic loan words which have become English repertoires (Brown, 1996), translated Arabic-origin words, transliterated Arabic origin words (Hassan, 2016), or a combination of transliterated Arabic words accompanied by a description in English.

Yet, studies in this topic suggest ambiguous stands to the use of Islamic-oriented lexical items. For example Erlina et al., (2016 ) reported that the Arabic-originated word "*surah*" was

used in its original lexis instead of presenting their one-to-one counterpart or any similar vocabulary in English (p. 51). In the meantime, Hassan (2016) recommended transliteration for such lexical items that are highly cultural or ideological, hence have “partial-equivalent” or even “non-equivalent” (p.126) counterparts in English. On the other hand, Brown (1996a) Brown, (1996, p. 2) pointed out that “It may surprise some readers that many Arabic Islamic words are in English dictionaries.” For example ‘ayatollah’ appears in Longman Dictionary of Contemporary English, Oxford Advanced Learners’ Dictionary, Collins COBUILD English Dictionary (COBUILD) and Cambridge International Dictionary of English.

The other point of concern, regarding the choice of using locally developed materials (textbooks and course books), is the issue of mismatch of the texts and tasks presented in these materials with the “real language”. The findings of currently available studies pointed out that the writers of ELT textbooks/coursebooks often relied on their intuitions rather than consulting empirically developed language corpora nor word lists when selecting the language items to be presented in the textbooks/coursebooks (Alavi & Rajabpoor, 2015; Grammatosi & Harwood, 2014). One of the practices of intuition-based materials was exemplified by the re-writing of reading texts in the textbooks into simpler versions after the original texts are collected (Chen, 2016).

Acknowledging the importance of corpora and word lists in English for Specific Purposes (henceforth ESP), a number of studies have been conducted and focused primarily on producing technical word lists in various disciplines. Some of the examples of those specialized vocabulary lists include a word list for finance (Kwary, 2011), pharmacy (Grabowski, 2013, 2015), nursing (Mohamad & Ng, 2013; Yang, 2015), agriculture (Muñoz, 2015), newspaper (Zhu, 2017), social sciences (Kwary & Artha, 2017), and plumbing (Coxhead & Demecheleer, 2018).

The pedagogical usefulness of specific purposes corpora and word lists developed on the basis of the occurrences for certain vocabulary items in specific texts has been related to the argument that frequency information is important in language acquisition (Davies & Gardner, 2010; Gablasova, Brezina, & McEnery, 2017; Lei & Liu, 2016). Experts argued that frequency and range of the lexical items are the key information for vocabulary selection in ELT textbooks (Flowerdew, 2012; Grammatosi & Harwood, 2014; Mukundan & Kalajahi, 2016; Römer, 2010). For a word list to be claimed pedagogically useful, it has to be designed to address the

complexity and dynamic process of language learning. This means that a “one for all” word list is impossible and that specific type of learners required a specific ‘type’ of word list (Brezina & Gablasova, 2017). Given their unique needs to master English alongside Arabic (the most important language in Islamic studies), the learners of English in Indonesian Islamic universities are in need of a unique word list containing important vocabulary to meet these needs. This research, therefore, strives to find important vocabulary in the area of Islamic religious studies (henceforth IRS) using a corpus-based approach through the development of a word list related to Islamic studies.

## **1.2. Research Questions**

A language corpus is prescribed to be developed by adhering to certain guidelines related to the purposes, method, and content. Similar principles are set for word lists, including technical vocabulary lists (Flowerdew, 2012; Nation, 2016; Reppen, 2009b). Such principles must be maximally followed in order for corpora and word lists to comprehensively describe the features of the languages they represent, and also to have meaningful pedagogical impact on the teaching of those languages. The research questions addressed in this study are set surrounding the issue of providing a specifically developed target corpus and technical word/vocabulary list that addressed the needs of the target ELT context whilst fulfilling these guidelines. Hence, the research questions crafted in this present study are as follows:

1. What lexical items or word-types occur frequently and uniformly across the textbooks of Islamic studies taught at an Indonesian State Institute for Islamic Studies (IAIN/ISIIS)?
2. Based on the corpus analysis of the IRS textbooks taught at the ISIIS/IAIN, which word-types should be included in the Islamic Religious Studies Technical Vocabulary (IRSTV) list?
3. How is the application of the IRSTV to inform ELT vocabulary materials for IAIN/ISIIS students?
  - a. What words types or vocabulary items are to be included in ELT materials for IAIN/ISIIS students? And,

b. How to use the IRSTV list as the language content of ELT vocabulary materials to meet the representativeness of the English necessary for teaching academic/ESP in the field of Islamic studies?

### **1.3. Research Objectives**

In line with the stated research questions, the study embarks on the following objectives:

1. To describe the categories of vocabulary that build the CIRELISST-ISIIS i.e. by profiling the vocabulary content based on Nation's (2001) vocabulary classification concept.
2. To create a technical word list that is specifically tailored to represent English vocabulary (i.e. word-types) in the discipline of Islamic religious studies (IRS).
3. To demonstrate the application of the IRSTV list in corpus-informed materials development/design.

### **1.4. Scope of the Study**

The present study focuses on creating an English technical vocabulary list in the discipline of Islamic Religious Studies (IRS) using quantitative measures of frequency-range and keyness scores, which are combined into the RFK score. The scope of this present study is specified as a corpus-based technical word list development based on the target corpus i.e. the CIRELISST-ISIIS is written academic field-specific English contained in the obligatory textbooks or references of the five main subjects of IRS taught at the examined IAIN/ISIIS. Other measures for constructing field-specific word lists including a mixed-method analysis combining quantitative and qualitative criteria is not employed, although this analysis is seen as more comprehensive than if each of these criteria is used in isolation. The size of the target corpus is also limited to only approximately 305,000 running words/tokens. This small size is still theoretically acceptable as the scope for a specialized corpus like the present study's CIRELISST-ISIIS.

## 1.5. Significance of the Study

The present study of EFL materials development put forward in this study is geared toward accommodating the corpora of English as used in the area of religion or faith-related corpora. This covers general academic and professional corpora related to Islamic studies and or English as used in countries whose Muslim population speak ESL/EFL even English as a first language (see Al-Ali, 2010; Alcoberes, 2016).

This present study is argued to be a significant one, based on these following points: first, in terms of theoretical significance, this present study is to contribute to the enrichment of literature of materials development in EFL context addressing the less explored topics of research of the evaluation of the lexical items of ELT textbooks/course books from the perspective of corpus approach in Indonesian context. The present study, more specifically, extends the work of Abudukeremu (2010) on the exploring lexical items that are used in Islamic academic research articles that resulted in the development of the corpus of Islamic academic research articles (IARA). The result of the present study is expected to complement Abudukeremu's (2010) work by exploring the more general occurrence of IRV in the core textbooks of IRS written in general academic English in Islamic studies disciplines.

This present study is built upon and continues previous studies in linguistics and applied research on English as used by Islamic academic and professional communities (Al-Ali, 2010; Al-Faruqi, 1986; Hassan, 2016; Jasem, 1995; Jassem, 2013; Mahboob, 2009). The present study also complements numerous researches on the content of ELT learning resources that examine the inclusion or exclusion of topics, language features, skills as well as some other educational or pedagogical methods. These studies include evaluation of quality the textbook/course books and or materials content against some sets of evaluative checklists examining the linkage of textbooks/course books/materials contents with general learning objectives as suggested by the curriculum. Some works on this topic that have flourished academic publication examine the content of textbooks based on the enclosure of certain educational or learning concepts sufficient coverage of language skills or tasks within a specified context of ELT (Rukmini, 2007; Widodo, 2015). Other studies that investigates locally developed textbooks for EFL learners address the locally written texts of these textbooks and their implications to EFL learning (Alcoberes, 2016; Mukundan & Kalajahi, 2013). In addition, there are issues such as equal treatment of language

and culture content, as well as representation of local versus the Western/Anglo-Saxon culture associated with English (Azizifar & Baghelani, 2014a; Lawrence, 2011). Another popular issue is the appropriateness of the content of internationally published textbooks against the local cultural and ideological values of their users (Ashraf, Motallebzadeh, & Kafi, 2013; Stranger-Johannessen, 2015).

The second point of significance is a practical one, in that it is related to the existing materials/textbooks/course books used by certain EFL programs in Indonesian Islamic universities. These textbooks are either form-based ones, for example, Kardimin (2011) or skill-based, for example, Arsyad (2001). The integration of the “Islamic” aspect of Indonesian Islamic higher education in these textbooks is by presenting texts related to the ‘general Islamic teaching’ enveloped with the teaching of EFL as desired by the context. The present study offers NNS of English teachers and students some insights into choosing more suitable teaching and learning materials in Islamic academic context based on a careful implementation of research findings into materials development. This includes the studies highlighting the importance of linking ELT textbooks’ language with “the real language” as shown by corpora of English.

The triangulated range-frequency-keyword (RFK) analysis based on Corpus Linguistics approach provides strength to the present ELT materials by providing empirically supported vocabulary in Islamic studies. This feature is absent in other vocabulary materials designed for Muslim learners of ESL/EFL, due to the intuitive nature of the development and selection of IRV. The frequency and keyword result that serves as the basis of sequencing the IRV in the present ELT materials in this research are other advantages over other IRV lists which is not informed by these methods. The IRVs in this kind of encyclopedia, therefore, need further treatment in order to be pedagogically-ready. The IRVs resulted from this current research, on the other hand, can be sequenced on the basis of their occurrence in “the real language” of the corpus of Islamic religious studies core textbooks used in Indonesian Islamic universities, as suggested by the theory of vocabulary learning and developing vocabulary learning programs proposed by (P. Nation, 2001, 2016). In all, the currently present research project will contribute positively to ELT teaching in this specific context as well as offers some insights into the more particular theory and practice of ELT materials development.

## 1.6. Clarification of the Terms

The following is the clarification of the key terms of this present study, which is very often referred to in the entire report:

1. Islamic Religious Studies: An academic discipline that examines the teachings of Islam as an independent science, covering theology, law, philosophy, history, and creed. In an Indonesian context, Islamic religious studies uniquely combined Islamic intrinsic perspective with Western views (Azra, 2011, p. 44).
2. Technical vocabulary: vocabulary that has special meanings that are often unique to certain discipline, knowledge or vocational field. Sometimes technical words have general meanings which are understood in the common domain (P. Nation, 2001).
3. EFL learners: learners of English as a foreign language (EFL) who are studying at Indonesian Islamic universities or colleges.
4. Indonesian Islamic universities: Universities or colleges that are established to specifically study and research on the teachings of Islam as academic sciences. In Indonesia, these universities and colleges are managed under the Ministry of Religious Affairs (MORA), whether they are founded and funded by the Indonesian government or by private foundations.
5. Corpus: originally a Latin word that means ‘body’ is defined as “... a body of language representative of a particular variety of language or genre which is collected and stored in electronic form for analysis using concordance software” (The ESRC Centre for Corpus Approaches to Social Science (CASS), 2013, p. 5).
6. Corpus-based study: A study that uses corpus data and methods to approach a language or a linguistic approach which is conducted using corpus methods (The ESRC Centre for Corpus Approaches to Social Science CASS, 2013)
7. Materials development: Theoretically means a field of study that investigate the principles and procedures of the design, implementation, and evaluation of language teaching materials. Practically materials development refers to every activity of supplying information about and/or experience of the language in ways designed to promote language learning.

8. Word-types: one single 'kind' of a word, for example in the word string of five running words 'all of you except you' contains only four word-types. When the same word i.e. 'you' occurs again, its second occurrence is not counted again.
9. Corpus analysis: a linguistic analysis using corpus method aided by software designed specially for corpus analysis.

## **1.7. Organization of the Thesis**

This thesis is divided into five chapters as follows:

### **Chapter I: Introduction**

The first chapter consists of sub-chapters i.e. the background, research questions and objectives, the scope of the study, clarification of the terms, and research significances.

### **Chapter II: Theoretical Framework and Literature Review**

The second chapter contains the theoretical foundations of this present study, that is, the theory of Corpus Linguistics and ELT Materials Development. The sub-chapters discuss definitions, main principles, typology and previous studies that have been conducted in each of the two main theories.

### **Chapter III: Research Method**

This chapter reviews the research methodology that is used in conducting the present study. This includes elaboration on the research design and the methods of the construction of the target corpus and the IRSTV list. The detailed procedures of constructing each of the respective 'products' i.e. the CIRELISST-ISIIS and the IRSTV list is presented.

### **Chapter IV: The Corpus of Islamic Religious Studies Textbooks (CIRELISST) and the Islamic Religious Studies Textbooks Vocabulary (IRSTV) List.**

The chapter presented the result of the CIRELISST-ISIIS development, the IRSTV List created using the texts data in the CIRELISST-ISIIS and the more practical account on the application of the IRSTV List into text selection and vocabulary grading.

## Chapter V: Conclusions, Limitations of the Study and Recommendations

This concluding chapter is comprised of sub-chapters of the conclusions made based on the findings of the development of a specific target corpus i.e the CIRELISST-ISIIS, the creation of the IRSTV List and its application. The chapter also touches upon several issues deemed as limitations of this present study. It is also pointed out that such limitations are actually an open door for further research endeavors under the same interest with this present study. In addition to the more theoretical recommendations for future research projects, more practical suggestions on the actual pedagogical uses of the IRSTV List in tertiary level Islamic universities and colleges in Indonesia.

\*\*\*\*\*