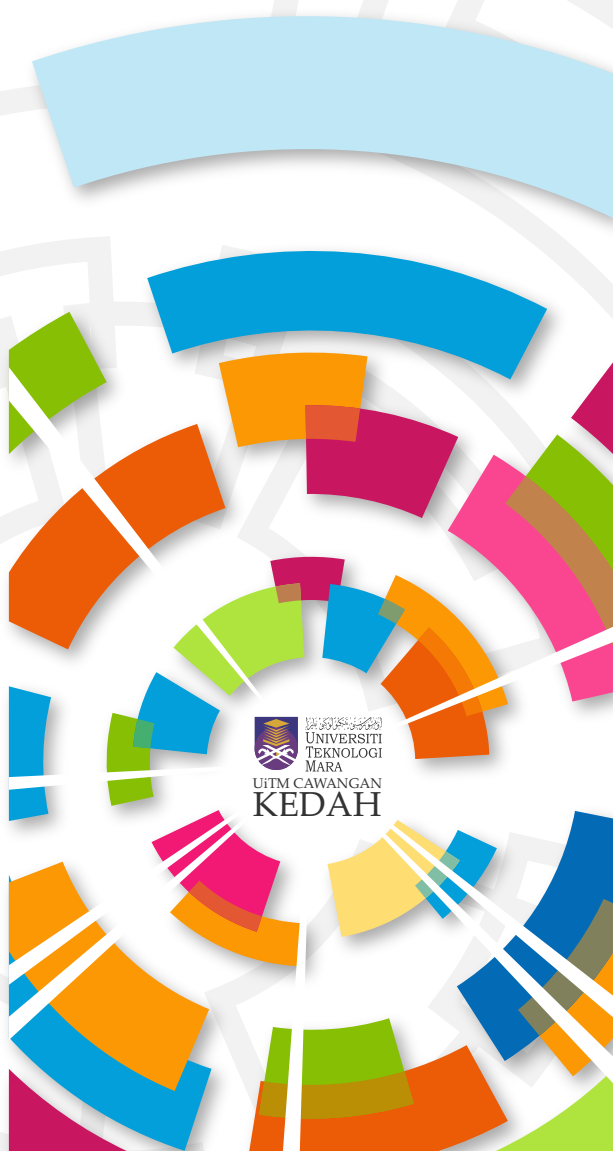


Voice of Academia

Academic Series of Universiti Teknologi MARA Cawangan Kedah

Vol. 9, No. 1, 2014 ISSN 1985 - 5079



ACADEMIC SERIES COMMITTEE UNIVERSITI TEKNOLOGI MARA CAWANGAN KEDAH

Editor-in-Chief

Dr Mahzir Ismail
Sharina Saad

Advisory Board

Aishah Musa
Etty Harniza Harun
Hasroleffendy Hassan
Nor Ardyanti Ahmad

Syakirah Mohamed
Syed Mohammed Alhady Syed Ahmad Alhady
Prof. Madya Wan Faizah Wan Abdullah

Graphic

Asrol Hasan

Special Tasks

Aishah Musa
Bawani Selvaraj
Ho Chui Chui

Rita Maria Dourado
Siti Ajar Ikhsan (USM)
Syazliiyati Ibrahim

Review Board

Prof Madya Dr. Janudin bin Sardi@Mohd Yusop
Universiti Teknologi MARA Shah Alam
Dr Rafizah Mohd Rawian
Universiti Teknologi MARA Perlis
Dr. Norizul Azida Darus
Universiti Teknologi MARA Perlis
Dr Kamisah Ariffin
Universiti Teknologi MARA Pahang,
Dr Irene Leong
UiTM Kampus Bandaraya Melaka
Dr Sujatha Menon
UiTM Kampus Bandaraya Melaka

Dr. Mohd Norawi Ali
Universiti Sains Malaysia
Dr Muhammad Saiful Haq bin Hussin
Universiti Malaya
Dr Muhammad Saiful Haq bin Hussin
Universiti Malaya
Dr Mohd Zaki Abd Rahman
Universiti Malaya
Dr Ghazali Bin Lateh
Universiti Putra Malaysia

Copyright © 2014 by the Universiti Teknologi MARA, Kedah

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or any means, electronic, mechanical, photocopying, recording or otherwise, without prior permission, in writing, from the publisher.

© *Voice of Academia* is jointly published by the Universiti Teknologi MARA, Kedah and Penerbit UiTM (UiTM Press), Universiti Teknologi MARA Malaysia, Shah Alam, Selangor. Printed by Perpustakaan Sultan Badlishah, Universiti Teknologi MARA, Kedah.

The views, opinions and technical recommendations expressed by the contributors and authors are entirely their own and do not necessarily reflect the views of the editors, the Faculty or the University.



Making A Placement Test Based On Vocabulary Knowledge

*Andrew Gorringe
John Anderson*

Malaysian Undergraduates' (In)ability to Identify the Correct Use of Sources

*SHo Chui Chui
Wong Soon Heng*

Keberkesanan Pembelajaran Kemahiran Lisan Berstruktur Dan Penggunaan ICT Bagi Penilaian Lisan

Roslinda Johari

Pembinaan Data Korpus Bahasa Arab Untuk Tujuan Pengajaran Di Peringkat STAM

*Zainur Rijal B. Abdul Razak (Dr)
Rosni B. Samah
Muhammad Fauzi B. Jumingan
Mohd. Shahrیمان B. Abu Bakar*

Program Literasi Bahasa Melayu Pelajar Tahun Tiga Sekolah Kebangsaan Tiong, Kota Bharu Kelantan Bekerja di Tanah Air

*Helmey Zainee Binti Mohd Zain
Zurairhan binti Zakaria*

CONTENTS	PAGE
Making A Placement Test Based On Vocabulary Knowledge	1
<i>Malaysian Undergraduates' (In)ability to Identify the Correct Use of Sources</i>	13
Keberkesanan Pembelajaran Kemahiran Lisan Berstruktur Dan Penggunaan ICT Bagi Penilaian Lisan	23
Pembinaan Data Korpus Bahasa Arab Untuk Tujuan Pengajaran Di Peringkat STAM	33
Program Literasi Bahasa Melayu Pelajar Tahun Tiga Sekolah Kebangsaan Tiong, Kota Bharu Kelantan Bekerja di Tanah Air	55

Pembinaan Data Korpus Bahasa Arab Untuk Tujuan Pengajaran Di Peringkat STAM

Zainur Rijal B. Abdul Razak (Ketua Penulis)
Fakulti Pengajian Bahasa Utama
Universiti Sains Islam Malaysia
Bandar Baru Nilai, 71800, Negeri Sembilan
Tel: +60196641102 E-mail: zrijalrazak@gmail.com

Rosni B. Samah
Fakulti Pengajian Bahasa Utama
Universiti Sains Islam Malaysia
Bandar Baru Nilai, 71800, Negeri Sembilan
Tel: +60196694261 E-mail: rosni@usim.edu.my

Muhammad Fauzi B. Jumingan
Fakulti Bahasa Moden Dan Komunikasi
Universiti Putra Malaysia
Serdang, 43400, Selangor
Tel: +603-8946 8794 E-mail: uzi@fbmk.upm.edu.my

Mohd. Shahrman B. Abu Bakar
Pusat Pengajian Bahasa Dan Pembangunan Insaniah
Universiti Malaysia Kelantan
Jalan Pengkalan Chepa, Kota bahru, Kelantan
Tel: +6012 3766958 E-mail: shahrman@umk.edu.my

ABSTRAK

Beberapa kajian dalam bidang linguistik korpus telah membuktikan bahawa data korpus memainkan peranan besar dalam membantu pelajar memahami subjek yang dipelajari. Ini kerana ia dapat memberi maklumat tentang kosa kata dan kata kunci penting dalam subjek tersebut. Tujuan kertas kerja ini disediakan, pertama, adalah untuk menghuraikan langkah yang telah dilalui dalam membina data korpus bagi dua buku teks di peringkat STAM bagi subjek Fiqh dan Hadith. Kedua, mengenal pasti kosa kata penting dan kata kunci dalam dua buku teks tersebut. Bagi mencapai objektif di atas, data yang mengandungi sekitar 124,500 perkataan dipindahkan dalam bentuk softcopy dan dilakukan proses pengkodan. Kemudian data tersebut dianalisis menggunakan perisian Wordsmith 6.0 untuk mendapatkan frekuensi penggunaan perkataan dan kata kunci bagi setiap subjek. Dapatan menunjukkan bahawa satu korpus lebih besar yang

merangkumi semua subjek di peringkat STAM mampu dibina. Analisis ke atas korpus mendapati, terdapat beberapa kosa kata penting yang umum digunakan dan kata kunci bagi setiap subjek.

Kata Kunci: Pembinaan korpus, buku STAM, analisis korpus, frekuensi perkataan, kata kunci

1. Pengenalan

Memiliki perbendaharaan kata yang luas dalam mempelajari sesuatu ilmu dalam bahasa asing merupakan salah satu faktor penting kepada kefahaman menyeluruh terhadap ilmu yang dipelajari. Ini kerana ia dapat membantu pelajar memahami subjek-subjek yang dipelajari dalam bahasa asing itu tidak kira sama ada di peringkat sekolah mahupun universiti. Menurut Richard Anderson dan William Nagy (1984), bagi memahami subjek-subjek yang dipelajari di peringkat sekolah menengah, seseorang pelajar itu mesti mampu memahami antara 2,000 hingga 3,000 patah perkataan umum baru setiap tahun. Sementara itu, kajian Liu Na dan Paul Nation (1985) pula mendapati sebanyak 3000 perkataan asas mesti dikuasai oleh pelajar di peringkat universiti dalam bidang tertentu.

Permasalahan yang timbul pada masa kini khususnya di peringkat Sijil Tinggi Agama Malaysia (STAM) adalah para pelajar tidak mengetahui apakah kosa kata yang sepatutnya diberi keutamaan untuk mereka fahami. Sebahagian besar pelajar sekadar menghafal makna sejumlah perkataan tanpa mengetahui tahap kepentingannya dalam sesuatu subjek, sebaliknya perkataan yang tidak diberi keutamaan pula amat penting dan menjadi kata kunci bagi subjek tersebut. Justeru, pembinaan korpus berdasarkan buku-buku teks yang digunakan dalam aktiviti P&P pelajar di peringkat ini dilihat mampu memberikan gambaran jelas tentang apakah kosa kata yang mesti diberi keutamaan dan sebaliknya. Bukan sekadar itu, proses pembelajaran akan menjadi lebih pantas dan sistematik jika pelajar dapat mengaitkan pula antara setiap kosa kata penting yang menjadi kata kunci mengikut subjek dengan topik yang dipelajari.

2. Pernyataan Masalah

Lazimnya, penggunaan sesuatu kosa kata di dalam buku teks ditentukan oleh bidang dan peringkat pengajian itu sendiri. Semakin tinggi peringkat

pengajian dan semakin khusus bidang yang dipelajari semakin sukar perkataan yang digunakan dan semakin banyak pula kosa kata yang perlu dikuasai. Pengajian di peringkat STAM di sekolah menengah yang menggunakan bahasa Arab sebagai bahasa pengantar dianggap amat penting kerana ia menjadi satu bidang khusus kepada pelajar-pelajar jurusan agama serta menjadi penentu hala tuju mereka meneruskan pengajian di peringkat yang lebih tinggi. Namun begitu, hasil kajian lepas mendapati bahawa pelajar dalam bidang ini menghadapi kesukaran untuk mengikuti pengajian mereka di universiti (Nur Zafirah, 2011).

Antara punca masalah ini adalah kurangnya perbendaharaan kata dalam kalangan mereka semenjak di peringkat sekolah menengah lagi. Dengan membuat penentuan secara pasti dan konkrit tentang apakah kosa kata yang mesti dikuasai oleh pelajar di peringkat STAM melalui pembinaan data korpus, masalah kelemahan mereka di peringkat sekolah dan universiti diharap dapat diselesaikan. Ini bertepatan dengan pandangan Tony McEnery dan Richards Xiao (2011), yang menyatakan penggunaan korpus dalam P&P amat signifikan kerana ia menentukan apakah bahan pengajaran yang perlu diajar kepada pelajar dan bagaimana bahan itu diguna pakai.

Justeru, permasalahan kajian ini adalah bertitik tolak daripada kesukaran pelajar menguasai bahasa Arab dan subjek-subjek agama di peringkat STAM. Masalah ini mengakibatkan kesan lebih parah apabila mereka melanjutkan pengajian di peringkat universiti yang banyak bergantung kepada kaedah pembelajaran sendiri. Salah satu punca masalah yang dapat dikenal pasti adalah ketidakfahaman mereka terhadap sebahagian kosa kata yang digunakan dalam buku-buku teks serta ketidakmampuan mereka mengenal pasti apakah kata kunci dalam subjek tersebut. Akibat masalah ini juga, dalam sesetengah keadaan, para pelajar tidak menyedari bahawa terdapat sebahagian kosa kata bersifat ulangan di dalam buku teks yang digunakan. Gabungan masalah-masalah ini secara tidak langsung telah melambatkan proses pemahaman isi kandungan subjek itu dan merencatkan pencapaian akademik mereka.

Permasalahan ini dapat diatasi apabila satu korpus khas di peringkat STAM yang membolehkan pelajar mengetahui apakah perkataan-perkataan yang kerap digunakan dalam buku-buku teks dan kata kunci bagi setiap subjek dapat dibina. Keadaan ini amat mendesak pada masa kini memandangkan korpus yang mengumpulkan data dari buku-buku teks di peringkat STAM

masih belum pernah dibangun oleh mana-mana pihak.

Dengan menganalisis korpus yang dibina, para pelajar dapat mengetahui apakah kosa kata penting mengikut subjek yang perlu difahami terlebih dahulu. Melalui pemahaman terhadap kosa kata penting dalam subjek yang dipelajari, proses pembelajaran akan menjadi lebih pantas dan sistematik.

3. Kajian Literatur

Melihat dari aspek sejarah perkembangan, usaha membina data korpus telah bermula semenjak akhir kurun ke-19 lagi di negara barat tetapi korpus pada ketika itu hanya merupakan sejumlah teks yang terhad dalam bentuk fizikal (John T. Fotos, 1931). Perkembangan yang pesat dapat dilihat apabila komputer mula dicipta yang membolehkan teks-teks dikumpulkan dalam bentuk digital dan memudahkan proses mengkod data. Antara korpus terawal dalam bahasa Inggeris yang pernah dibangun adalah Brown corpus, Lancaster-Oslo-Bergen corpus (LOB) (Knut Hofland & Stig Jahansson: 1987), FLOB, FROWN, ACE dan Wellington NZ Corpus (al-Sulaiti & Atwell: 2006). Tidak ketinggalan juga korpus yang dibina dalam bahasa-bahasa lain seperti bahasa Jerman, Sepanyol, Itali, Perancis dan sebagainya. Semua korpus ini mengandungi sejumlah besar teks dalam bahasa-bahasa tertentu dan sedia dianalisis dari pelbagai aspek bahasa seperti morfologi, sintaksis, gaya bahasa dan sebagainya.

Hal ini tidak kurang hebatnya dalam bahasa Arab walaupun agak terlewat bermula berbanding bahasa lain. Jadual di bawah menunjukkan beberapa contoh korpus dalam bahasa Arab yang pernah dibina (Latifa al-Sulaiti & Eric Atwell: 2006).

Jadual 1: Contoh korpus bahasa Arab yang pernah dibina

Nama Korpus	Medium	Jumlah Perkataan	Tujuan	Data
Buckwalter Arabic Corpus (1986-2003)	Penulisan	2.5 – 3 juta	Perkamusan	Data umum laman web
Leuven Corpus (1990-2004)	Penulisan dan pertuturan	3 juta (700,000 pertuturan)	Perkamusan	Internet, radio, televesyen dan buku sekolah rendah
Arabic Newswire Corpus (1994)	Penulisan	80 juta	Pendidikan pengembangan teknologi pendidikan	AFP, Agensi Berita Xinhua dan Berita Umma
CALLFRIEND Corpus (1995)	Perbualan	60 perbualan telefon	Pembangunan teknologi bahasa	Penutur jati Mesir
CLARA (1997)	Penulisan	50 juta	Perkamusan	Jurnal, buku, dan Internet
An-Nahar Corpus (2001)	Penulisan	140 juta	Penyelidikan	Surat khabar an-Nahar, Lubnan
Al-Hayat Corpus (2002)	Penulisan	18.6 juta	Kejuteraan bahasa	Surat khabar Al-Hayat, Lubnan
Classical Arabic Corpus (2004)	Penulisan	1.6 juta	Analisis perkamusan	Internet

Sumber: http://www.comp.leeds.ac.uk/eric/latifa/arabic_corpora.htm

Dalam pembinaan korpus lazimnya, para sarjana mengkategorikan data yang dikumpul mengikut kategori teks dan dinamakan sebagai subkorpus, contohnya korpus Al-Hayat mengkategorikan teks kepada teks umum, kereta, komputer, berita, ekonomi, sains dan sukan, manakala korpus An-Nahar membahagikan kepada kategori umum, politik, berita, komputer, sastera, ekonomi dan sukan (ELDA Home page, t.t.). Dalam korpus STAM ini, pengkaji telah membahagikan subkorpus mengikut kategori subjek iaitu Fiqh dan Hadith.

Idea pembinaan korpus pada asalnya adalah sebagai data bagi pembinaan kamus. Kemudian peranannya berkembang kepada aplikasi data korpus di dalam P&P di pelbagai peringkat pengajian. Lanjutan daripada itu, penggunaannya dalam P&P telah menghasilkan banyak teori dan dapatan yang saintifik terutama dalam meningkatkan kemampuan bahasa pelajar. Menurut pandangan Carmen Llantada (2009), penggunaan data korpus dalam pembelajaran kemahiran bertutur dalam bahasa Inggeris memberi kesan positif terhadap kemahiran pelajar terutama dari aspek tatabahasa. Li-szu Agnes (2012) telah menjalankan kajian terhadap keberkesanan penggunaan data korpus dalam pembelajaran bahasa dan mendapati bahawa data korpus amat berguna dalam pembelajaran pelajar tidak kira sama ada dalam pembelajarannya sebagai bahasa pertama atau kedua.

Pada tahun 1982, Mustafa Ruslan telah menjalankan kajian korpus ke atas 65,966 perkataan dari penulisan kefahaman 120 orang pelajar tahun 6 di sekolah rendah di Mesir. Beliau mendapati pelajar di bandar memiliki lebih banyak perbendaharaan kata iaitu sebanyak 35,298 perkataan berbanding pelajar di kampung dengan hanya 30,654 perkataan. Sementara Paul Nation dan Robert Waring (t.t.), menyatakan bahawa seseorang pelajar yang menguasai 1000 perkataan, ia secara amnya mampu memahami 72% dari sesuatu teks yang dibaca, 2000 perkataan dapat memahami sebanyak 79.7% dari keseluruhan teks dan sehinggalah kepada 15,851 perkataan yang meliputi 97.8% teks. Secara keseluruhannya, mereka menyatakan bahawa penguasaan sebanyak 3000-5000 perkataan sudah memadai untuk pemahaman asas sesuatu teks yang dibaca.

Menurut kajian Xue Gouyi dan Paul Nation (1984), pelajar perlu menguasai lebih kurang 830 perkataan dalam sesuatu bidang subjek di peringkat universiti. Ini bermaksud, sebagai contohnya di bidang pengajian Islam, pelajar perlu memahami 830 perkataan dalam subjek bahasa Arab dan 830

perkataan lagi berkaitan al-Quran atau hadith dan sebagainya. Hasil dari perbincangan ini dapat disimpulkan bahawa pembinaan korpus dan analisis ke atasnya dapat memberikan maklumat yang tepat tentang kosa kata penting dalam sesuatu subjek yang mereka pelajari dan seterusnya dapat meningkatkan kefahaman mereka.

4. Kepentingan Korpus di Peringkat STAM

Umumnya, penguasaan pelajar terhadap kosa kata penting dan kerap digunakan di peringkat STAM banyak membantu mereka menguasai subjek-subjek yang dipelajari. Dalam proses mengenal pasti dan memilih perkataan-perkataan penting dan sering digunakan di peringkat pengajian mereka, buku-buku teks yang menjadi rujukan utama dalam proses pengajaran dan pembelajaran boleh dijadikan sandaran utama.

Kepentingan membina korpus khas di peringkat STAM dapat dilihat melalui kemampuannya mengumpulkan semua teks yang terdapat di dalam buku-buku teks STAM bagi membolehkan analisis dilakukan terhadap frekuensi penggunaan sesuatu perkataan dan disusun mengikut susunan (ranking). Susunan perkataan bermula dari yang paling kerap hingga ke senarai yang paling jarang digunakan. Oleh itu, pelajar dapat mengetahui apakah kosa kata penting yang kerap digunakan dalam buku-buku teks mereka dan kata kunci bagi sesuatu subjek itu. Kosa kata yang paling tinggi frekuensinya merupakan kosa kata yang paling kerap digunakan dan penting untuk dikuasai. Selain daripada itu, dengan membandingkan antara data mengikut subjek, memberikan gambaran kepada pelajar tentang kata kunci bagi subjek yang dianalisis.

Oleh itu, dapatan kajian ini juga diharap dapat menghasilkan senarai kosa kata penting yang mesti dikuasai pelajar di peringkat STAM khususnya dalam dua subjek yang dikaji. Lanjutan dari hasil kajian ini juga satu perisian korpus di peringkat STAM akan dibina yang boleh dimuat turun di makmal-makmal komputer di setiap sekolah menengah bagi pelajar mempelajari dan membuat kajian tentang penggunaan bahasa Arab di peringkat STAM bukan sahaja untuk mengetahui aspek kosa kata dan kata kunci tetapi juga aspek morfologi, sintaksis, gaya bahasa dan sebagainya.

5. Objektif

Kajian ini dijalankan bagi mencapai objektif di bawah:

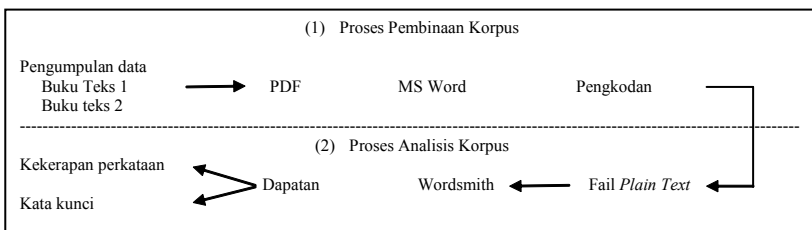
1. Membina korpus bagi dua subjek yang dikaji diperingkat STAM melalui data dari buku-buku teks STAM.
2. Mengenal pasti kosa kata yang kerap digunakan di dalam dua buku teks di peringkat STAM dan kata kunci-kata kunci bagi subjek-subjek tertentu di peringkat ini.

6. Metodologi

Oleh kerana semua buku teks STAM hanya diperolehi dalam bentuk hard-copy, dua buah buku teks yang dipilih diimbas dan dipindahkan dalam bentuk softcopy. Dua buku ini merangkumi dua subjek di peringkat STAM iaitu subjek Fiqh dan Hadith. Secara keseluruhannya korpus ini mengandungi sekitar 124,500 perkataan. Data yang telah dipindahkan ke bentuk softcopy disunting semula agar bersesuaian dengan perisian komputer yang akan digunakan dalam analisis frekuensi perkataan dan kata kunci.

Analisis frekuensi perkataan dan kata kunci dilakukan dengan menggunakan perisian Wordsmith 6.0. Perisian ini mempunyai tiga aplikasi utama bagi analisis statistik bahasa iaitu senarai perkataan, konkordan dan kata kunci (Micheal Wilkinson, 2012). Aplikasi ‘senarai perkataan’ digunakan untuk menyenaraikan semua perkataan yang terdapat dalam korpus mengikut susunan frekuensi. Manakala aplikasi ‘kata kunci’ digunakan untuk mencari kata kunci bagi setiap subkorpus dan membandingkan sesama subjek menggunakan ujian likelihood-ratio. Secara ringkas kajian ini melibatkan proses seperti di dalam carta alair di bawah:

Carta alir 1: Proses pembinaan dan analisis korpus



7. Pembinaan Korpus

7.1 Kandungan Data

Korpus yang dibina meliputi data dari dua buah buku teks bagi subjek Fiqh dan Hadith. Tajuk dua buku tersebut dan jumlah perkataan bagi setiap buku adalah seperti di jadual 2 di bawah:

Jadual 2: Tajuk buku teks dan jumlah perkataan

Subjek	Tajuk Buku	Jumlah Perkataan
Fiqh	<i>Al-Iqna'</i>	93,508
Hadith	<i>Al-Ahadith al-Mukhtarah min Fath al-Mubdi</i>	31,049
	Jumlah	124,557

Perbezaan jumlah perkataan yang jelas antara kedua buku teks tidak menjejaskan analisis perbandingan kerana ia telah disetarakan dengan menggunakan ujian likelihood-ratio yang terdapat di dalam perisian Wordsmith yang dipilih.

7.2 Langkah Pembinaan Korpus

Pembinaan korpus dalam kajian ini melibatkan 3 langkah:

7.2.1 Mengumpul data

Bagi membolehkan analisis korpus dilakukan secara saintifik menggunakan perisian komputer, pengkaji mesti memastikan data disediakan dalam bentuk salinan lembut (softcopy). Oleh kerana pengkaji hanya mempunyai data buku teks STAM dalam bentuk salinan keras (hardcopy), langkah ini melibatkan proses memindahkan teks-teks tersebut ke dalam bentuk softcopy. Semasa proses pemindahan tersebut, banyak kesalahan ejaan berlaku dan ini menuntut pengkaji melakukan proses suntingan berpandukan kepada teks asal.

7.2.2 Mengkod data

Sejauh mana proses mengkod data diperlukan adalah ditentukan oleh tujuan dan kemampuan perisian komputer yang akan menganalisis korpus yang

dibina. Ia penting dijelaskan supaya data yang dimasukkan ke dalam perisian komputer dapat dianalisis dengan tepat berdasarkan ciri bahasa yang dikehendaki pengkaji. Bertitik tolak dari hal ini, pengkodan data hanya melibatkan proses mengasingkan tambahan huruf di hadapan perkataan.

Dalam membincangkan hal ini, bahasa Arab secara umumnya merupakan bahasa yang kompleks dari pelbagai aspek bermula dari sebutan hurufnya hingga kepada aspek tatabahasa, sastera dan ilmu linguistiknya. Antara ciri bahasa Arab ialah kewujudan sambungan huruf tambahan di hadapan dan belakang perkataan seperti *ال المعرفة*, huruf *ل, ك, ي* dan sebagainya. Semua huruf bersambung ini akan merubah bentuk fizikal sesuatu perkataan itu. Memandangkan perisian yang digunakan menganalisis sesuatu perkataan mengikut sela di antara dua ruang kosong, maka proses mengkod dilakukan dengan memisahkan huruf sambungan ini dari perkataannya. Contohnya perkataan *الحديث* akan menjadi *ال حديث* dan perkataan *فيكتب* akan menjadi *في يكتب* setelah dipisahkan. Ini dilakukan supaya perisian tersebut dapat menganggap perkataan *الحديث* dan *حديث* sebagai satu perkataan yang sama dan dianalisis dalam satu entri.

Tanpa proses ini, kecatatan analisis akan berlaku kerana perkataan *الحديث* dan *حديث* serta perkataan *فيكتب* dan *يكتب* akan dianalisis sebagai dua perkataan berbeza akibat dari bentuk fizikalnya yang berbeza. Ini bertentangan dengan apa yang dikehendaki pengkaji. Namun, kewujudan fungsi lain dalam perisian Wordsmith yang mampu menyenaraikan perkataan mengikut susunan abjad mengurangkan beban pengkaji untuk memisahkan semua karektor perkataan kepada jujuknya. Justeru, hanya tambahan di bahagian hadapan perkataan sahaja dipisahkan. Sementara tambahan pada bahagian belakang seperti perkataan *هم* pada perkataan *كتابهم* tidak dipisahkan kerana keputusan susunan dapat juga dipaparkan mengikut abjad selain frekuensi.

7.2.3. Menyimpan data

Data yang telah dikod disimpan di dalam MS Word di dalam dua fail, pertama, fail yang diberikan tajuk *al-Iqna'* dan kedua, bertajuk *al-Ahadith al-Muthtarah fi Fath al-Mubdi*. Oleh kerana perisian Wordsmith yang digunakan dalam kajian ini tidak dapat memproses data dalam bentuk Word (*.docx), satu salinan data dipindahkan dalam bentuk plain text (*.txt)

dengan tajuk fail yang sama.

7.3 Pengecualian Perkataan

Seperti kelaziman dalam kajian teks, tidak semua perkataan dianalisis. Lebih-lebih lagi jika sesuatu kajian itu melibatkan jumlah perkataan yang sangat besar. Justeru, hanya perkataan yang membawa erti signifikan kepada objektif kajian sahaja akan dianalisis. Dalam kajian ini, perkara yang menjadi tumpuan adalah kosa kata penting dan khusus dalam dua buku teks yang dikaji seperti yang telah dijelaskan.

Oleh itu, perkataan yang dianggap terlalu umum penggunaannya serta tidak dapat menggambarkan gaya penggunaan perkataan yang signifikan telah dikecualikan. Ini termasuklah beberapa kategori perkataan seperti penanda wacana dan kata penghubung. Kategori perkataan ini memang banyak terdapat dalam apa jua jenis buku dan tidak memberikan banyak indikasi kepada gaya penggunaan perkataan. Atas alasan ini, hanya kategori kata nama dan kata kerja diambil kira sementara huruf diabaikan.

8. Analisis dan Dapatan

Perbentangan dapatan dalam kajian ini dibahagikan kepada tiga komponen utama:

1. Kosa kata penting dalam buku teks 1 - al-Ahadith al-Muhktarah min Fath al-Mubdi
2. Kosa kata penting dalam buku teks 2 – al-Iqna'
3. Kosa kata khusus mengikut buku teks (kata kunci)

8.1 Kosa kata penting dalam Buku Teks 1 – al-Ahadith al-Muhktarah min Fath al-Mubdi

Dalam kajian ini perkataan yang paling kerap digunakan dianggap paling penting dan mesti difahami pelajar. Kepentingannya dalam sesuatu subjek itu digambarkan melalui kekerapan penggunaannya. Selain daripada itu, kata kunci bagi sesuatu subkorpus juga mesti difahami pelajar bagi penguasaan menyeluruh ke atas sesuatu subjek. Jadual 3 di bawah menunjukkan senarai 30 perkataan paling kerap digunakan mengikut susunan.

Jadual 3: Senarai frekuensi 30 perkataan teratas dalam buku teks 1

Bil	Perkataan	F	%	Bil	Perkataan	F	%	Bil	Perkataan	F	%
1	الله	914	2.36	11	حديث	98	0.25	21	رواية	60	0.15
2	قال	546	1.41	12	قول	92	0.24	22	كسر	55	0.14
3	صلى	269	0.69	13	فتح	87	0.22	23	مراد	53	0.14
4	سلم	253	0.65	14	نسخة	79	0.20	23	يقول	53	0.14
5	تعالى	189	0.49	15	قيل	77	0.20	25	أول	48	0.12
6	كان	188	0.49	16	ضم	73	0.19	26	جنة	48	0.12
7	رضي	156	0.40	17	سلام	68	0.18	27	رجل	45	0.12
8	رسول	139	0.36	18	أمر	65	0.17	28	مهملة	45	0.12
9	نبي	138	0.36	19	معنى	63	0.16	29	مثل	44	0.11
10	صلاة	104	0.27	20	يكون	61	0.16	30	نفسه	41	0.11

Jadual 3 menunjukkan senarai kekerapan 30 perkataan yang paling kerap digunakan dalam buku teks 1 (*al-Ahadith al-Muhktarah min Fath al-Mubdi*). Ia menunjukkan bahawa perkataan yang paling kerap diulang adalah **نبي صلى الله عليه وسلم** dengan perkataan **نبي** sebanyak 138 kali, **صلى** 268, **الله** 914 dan **سلم** 253. Perkataan **عليه** tidak dianalisis kerana ia dari kategori huruf yang diabaikan. Dari aspek kata kerja, kata kerja **قال** (*telah berkata*) merupakan perkataan yang paling kerap digunakan dengan 546 kali. Ini bertepatan dengan ciri hadith yang dipindahkan dari satu generasi ke satu generasi melalui pertuturan sebelum zaman pembukuannya. Variasi lain bagi kata kerja ini juga terdapat di dalam senarai iaitu **قول** (*pengucapan*), **يقول** (*berkata*) dan **قيل** (*telah dikatakan*) dengan 92, 72 dan 53 frekuensi. Sementara itu, kewujudan perkataan **فتح**, **ضم**, **كسر** dan **ميم** menunjukkan bahawa buku ini amat mengambil berat terhadap ketepatan sebutan perkataan terutama perkataan yang jarang ditemui. Ia bukan sahaja diletakkan baris secara fizikal tetapi dinyatakan jenis barisnya dalam bentuk perkataan.

Jadual 4: Senarai frekuensi kata nama selepas 30 perkataan pertama dalam buku teks 1

Bil	Perkataan	F	%	Bil	Perkataan	F	%	Bil	Perkataan	F	%
1	رحمة	39	0.10	11	شيء	32	0.08	21	ترك	27	0.07
2	ناس	38	0.10	12	دنيا	31	0.08	22	أهل	27	0.07
3	خير	36	0.09	13	اسم	30	0.08	23	أرض	26	0.07
4	مسلم	36	0.09	14	مسجد	30	0.08	24	صفة	26	0.07
5	معجزة	36	0.09	15	يوم	30	0.08	25	ميم	26	0.07
6	أحدكم	33	0.09	16	أكثر	29	0.07	26	نساء	26	0.07
7	حال	33	0.09	17	القران	29	0.07	27	نبوة	25	0.06
8	سنة	33	0.09	18	بنيت	28	0.07	28	راء	24	0.06
9	همزة	33	0.09	19	حق	28	0.07	29	مرأة	24	0.06
10	سكون	32	0.08	20	نصب	28	0.07	30	واحد	24	0.06

Jadual 4 menunjukkan senarai kekerapan kata nama selepas 30 perkataan pertama dalam buku teks 1. Ia menunjukkan bahawa perkataan yang membawa maksud baris dan kaedah menyebut huruf sesuatu perkataan seperti yang telah dibincangkan masih kerap digunakan. Ia diwakili oleh kekerapan penggunaan perkataan *معجزة*, *همزة*, *سكون* huruf *ميم* dan *راء*. Perkataan-perkataan ini muncul antara perkataan yang kerap digunakan kerana ia terdapat dalam hampir semua subtajuk berbanding kebanyakan kata nama lain yang digunakan mengikut topik perbincangan sesuatu hadith.

Jadual 5: Senarai kekerapan kata kerja selepas 30 perkataan pertama dalam buku teks 1

Bil	Perkataan	F	%	Bil	Perkataan	F	%	Bil	Perkataan	F	%
1	قالت	39	0.10	11	أجيب	21	0.05	21	يتكلم	14	0.04
2	جمع	34	0.09	12	يجوز	21	0.05	22	يدل	13	0.03
3	كانت	31	0.08	13	ذهب	19	0.05	23	يكن	13	0.03
4	ذكر	28	0.07	14	سمعت	19	0.05	24	جعل	12	0.03
5	قالوا	28	0.07	15	أنزل	18	0.05	25	كانوا	12	0.03
6	قلت	26	0.07	16	خرج	18	0.05	26	مات	12	0.03
7	يقال	26	0.07	17	يدخل	18	0.05	27	يجب	12	0.03
8	يحتمل	25	0.06	18	يقل	16	0.04	28	يحل	12	0.03
9	فعل	24	0.06	19	تكون	15	0.04	29	يرد	12	0.03
10	رضى	22	0.06	20	سمع	14	0.04	30	تعلم	11	0.03

Jadual 5 menunjukkan senarai kekerapan kata nama selepas 30 perkataan pertama dalam buku teks 1. Ia menunjukkan bahawa pada frekuensi di antara 39 bagi perkataan *قالت* (*telah berkata*) dan frekuensi 11 bagi

perkataan **تعلم** (*mengetahui*) merupakan perkataan yang dianggap umum dan rata-rata pelajar peringkat STAM memahaminya. Walaupun begitu, dalam proses P&P yang sedang berjalan, para guru masih perlu memastikan semua pelajar memahami maksud perkataan-perkataan tersebut. Selain daripada itu, kewujudan perkataan seperti **قالت**, **ذكر**, **قالوا**, **قلت**, **يقال** dan **ينكلم** yang membawa maksud ‘menyatakan sesuatu’, mengukuhkan dapatan sebelum ini bahawa topik perbincangan dalam pengajian hadith berkisar tentang ucapan Rasulullah yang disampaikan melalui perantaraan lidah. Sementara kewujudan perkataan **رضى** (*telah diredhoi*) dengan 22 frekuensi adalah disebabkan variasi ejaan yang digunakan pada huruf **ي** (*ya*). Dalam kebanyakan keadaan di dalam buku ini, perkataan **رضى** dieja menggunakan huruf **ي** dengan dua noktah di bawahnya tetapi terdapat 22 frekuensi dieja tanpa noktah. Frekuensi yang lebih banyak bagi perkataan ini menggunakan dua noktah terdapat dalam jadual 3 dengan 156 frekuensi.

Jadual 6: Senarai kekerapan 30 perkataan teratas dalam buku teks 2

Bil	Perkataan	F	%	Bil	Perkataan	F	%	Bil	Perkataan	F	%
1	الله	583	0.48	11	قتل	188	0.16	21	حد	146	0.12
2	قال	568	0.47	12	ولد	179	0.15	22	خير	146	0.12
3	كان	551	0.46	13	أول	173	0.14	23	عتق	143	0.12
4	تعالى	407	0.34	14	دية	172	0.14	24	قاله	140	0.12
5	قوله	337	0.28	15	تنبیه	166	0.14	25	بيمين	140	0.12
6	مال	233	0.19	16	يكون	162	0.13	26	اولى	139	0.12
7	سلم	200	0.17	17	خلاف	156	0.13	27	زوج	135	0.11
8	طلاق	200	0.17	18	يجب	151	0.13	28	حق	134	0.11
9	صلى	193	0.16	19	حكم	150	0.12	29	أصل	131	0.11
10	ذكر	192	0.16	20	قطع	149	0.12	30	مصنف	128	0.11

Jadual 6 menunjukkan senarai kekerapan 30 perkataan tertinggi di dalam senarai kekerapan bagi buku teks 2 (*al-Iqna'*). Jadual menunjukkan bahawa perkataan yang paling kerap diulang adalah **الله** dengan dengan 583 frekuensi atau 0.48% dari jumlah keseluruhan perkataan. Oleh kerana perbincangan buku ini berkaitan hukum-hakam, kewujudan kata kerja **قال** (*telah berkata*) menunjukkan bahawa setiap hukum yang dibincangkan banyak mengaitkan ayat al-Quran. Sementara perkataan-perkataan lain seperti **مال** (*harta*), **قتل** (*membunuh*), **دية** (*diat*), **ولد** (*anak lelaki*), **زوج** (*suami*) dan **يمين** (*sumpah*) menunjukkan tajuk-tajuk perbincangan dalam buku ini berkaitan jenayah, wasiat dan perkahwinan. Berbeza dengan buku teks 1,

perkataan-perkataan khusus yang merujuk kepada tajuk tertentu berada pada 30 perkataan tertinggi. Sedangkan perkataan khusus pada buku 1 banyak terdapat pada frekuensi yang lebih rendah. Ini adalah disebabkan tajuk-tajuk yang dibincangkan dalam buku teks 2 lebih besar berbanding buku 1. Oleh itu kekerapan penggunaan perkataan yang sama adalah lebih tinggi.

Jadual 7: Senarai kekerapan kata nama selepas 30 perkataan pertama dalam buku teks 2

Bil	Perkataan	F	%	Bil	Perkataan	F	%	Bil	Perkataan	F	%
1	شهادة	122	0.10	11	أهل	109	0.09	21	يوم	90	0.07
2	شرط	120	0.10	12	سيد	108	0.09	22	أربعة	89	0.07
3	ثاني	118	0.10	13	أحد	107	0.09	23	أمة	88	0.07
4	نكاح	117	0.10	14	حر	107	0.09	24	كفارة	87	0.07
5	آخر	115	0.10	15	رجل	103	0.09	25	نسب	87	0.07
6	عدم	115	0.10	16	واحد	101	0.08	26	نصف	87	0.07
7	ثلاثة	113	0.09	17	سواء	99	0.08	27	آية	82	0.07
8	شيء	112	0.08	18	إسلام	93	0.08	28	كلام	82	0.07
9	ظاهر	111	0.09	19	رفيق	93	0.08	29	قصاص	80	0.07
10	نفسه	111	0.09	20	وطء	92	0.08	30	أكثر	76	0.06

Jadual 7 di atas menunjukkan senarai kekerapan kata nama selepas 30 perkataan pertama dalam buku teks 2. Ia menunjukkan bahawa perkataan yang digunakan pada peringkat ini banyak menjurus kepada tajuk-tajuk yang dibincangkan. Namun begitu terdapat juga beberapa perkataan lebih umum dan tidak menjurus kepada mana-mana tajuk seperti أكثر (*lebih banyak*), شرط (*syarat*), ثاني (*kedua*), آخر (*yang lain*), ثلاثة (*tiga*), شيء (*sesuatu*) dan sebagainya. Perkataan-perkataan ini biasanya sudah difahami pelajar di peringkat STAM kerana ia banyak digunakan dalam pelbagai tajuk dan situasi.

Jadual 8: Senarai kekerapan kata kerja selepas 30 perkataan pertama dalam buku teks 2

Bil	Perkataan	F	%	Bil	Perkataan	F	%	Bil	Perkataan	F	%
1	مر	127	0.11	11	بثبت	74	0.06	21	يحل	50	0.04
2	كانت	123	0.10	12	يكن	71	0.06	22	وجب	49	0.04
3	يجوز	120	0.10	13	تقبل	70	0.06	23	أعتق	46	0.04
4	فيل	113	0.09	14	يقطع	68	0.06	24	يأتي	46	0.04
5	خرج	108	0.09	15	يحرم	61	0.05	25	سبق	44	0.04
6	رضي	105	0.09	16	رواه	60	0.05	26	يسقط	42	0.03
7	مات	97	0.08	17	حرم	59	0.05	27	تجب	41	0.03
8	يشترط	90	0.07	18	يقول	59	0.05	28	يكفي	40	0.03
9	يصح	90	0.07	19	يعتق	54	0.04	29	تقدم	39	0.03
10	ذكره	74	0.06	20	يقتل	54	0.04	30	صرح	39	0.03

Jadual 8 menunjukkan senarai kekerapan kata kerja selepas 30 perkataan pertama dalam buku teks 2. Ia menunjukkan bahawa kata kerja yang digunakan tidak merujuk kepada mana-mana tajuk kecuali pada perkataan *أعتق* (*telah menebus*), *يعتق* (*menebus*) dan *يقتل* (*membunuh*) yang merujuk kepada penebusan kesalahan dan pembunuhan dalam tajuk jenayah dalam Islam. Sementara perkataan lain seperti *يجوز* (*boleh*), *يحرم* (*haram*), *يحل* (*halal*), *وجب* (*telah wajib*), *تجب* (*wajib*), *يصح* (*sah*), *حرم* (*telah mengharamkan*) dan *يشترط* (*mensyaratkan*) lebih merujuk kepada perkataan berkaitan hukum sesuatu perkara. Ini menunjukkan bahawa perbincangan dalam buku-buku fiqh Islami menekankan kepada perkara hukum-hakam walau apa jua tajuk yang dibincangkan iaitu jenayah, perkahwinan dan wasiat dalam konteks buku teks 2 ini.

Jadual 9: Analisis perbandingan antara buku teks 1 dan 2

Bil	Al-Ahadith kepada Al-Iqna'				Al-Iqna' kepada al-Ahadith			
	Perkataan	F (Al-Ahadith)	F (Al-Iqna')	Kata Kunci	Perkataan	F (Al-Iqna')	F (Al-Ahadith)	Kata Kunci
1	الله	914	583	914.64	دية	172	0	96.20
2	قال	546	568	317.32	تنبيه	166	1	83.44
3	رسول	139	26	263.45	عق	143	0	79.97
4	صلى	269	193	240.08	مال	233	11	71.69
5	سلم	253	200	204.83	مصنف	128	0	71.58
6	نبي	138	59	182.26	أب	140	1	69.23
7	رضي	156	105	147.51	طلاق	200	9	63.04
8	رواية	60	5	136.96	حر	107	0	59.83
9	حديث	98	41	131.05	قطع	149	5	53.33
10	سلام	68	13	127.92	ولد	179	9	53.24
11	ضم	73	23	113.27	نكاح	117	2	50.76
12	صلاة	104	62	108.96	يشترط	90	0	50.32
13	فتح	87	51	92.36	شهادة	122	3	48.38
14	جنة	48	10	87.79	زوج	135	5	46.46
15	رحمة	39	4	85.73	يمين	140	6	45.17

Jadual 9 menunjukkan keputusan perbandingan antara buku teks 1 dan 2 dari dua perspektif, pertama, perbandingan buku 1 kepada buku 2 dan kedua, perbandingan buku 2 kepada buku 1. Dua analisis ini dijalankan serentak bagi menjelaskan perkataan yang menjadi kata kunci bagi setiap buku teks. Pada analisis pertama (buku 1 kepada buku 2), didapati bahawa penggunaan ungkapan *قال نبي/رسول الله صلى الله عليه وسلم* paling banyak digunakan dalam buku teks 1 dan menjadi kata kunci bagi buku tersebut. Begitu juga dengan perkataan *رواية* (riwayat) dan *حديث* (hadith) yang merujuk secara langsung kepada perbincangan aspek hadith. Berbanding buku teks 2, buku teks 1 lebih mementingkan sebutan sesuatu perkataan dengan tepat. Ini dibuktikan dengan kewujudan perkataan *ضم* dan *فتح* yang merujuk kepada baris bagi perkataan-perkataan tertentu.

Bagi analisis kedua (buku 2 kepada buku 1), didapati bahawa semua perkataan yang menjadi kata kunci merujuk secara khusus kepada tajuk-tajuk yang dibincangkan dalam buku teks 2. Ini tidak termasuk perkataan *مصنف* (penulis) dan *تنبيه* (perhatian). Penggunaan yang banyak pada

perkataan pertama adalah disebabkan buku ini telah diberikan penerangan oleh ulamak lain terhadap matan Abu Syuja'. Perkataan ini digunakan apabila penulis buku merujuk kepada pemilik asal matan sepanjang perbahasanya. Pada perkataan kedua, penulis menggunakannya pada setiap permulaan tajuk kecil baru bagi memberi peringatan tentang pentingnya tajuk yang sedang dibincangkan. Justeru, para guru perlu memberikan penjelasan yang secukupnya tentang maksud perkataan-perkataan ini kepada pelajar dalam proses P&P.

9. Kesimpulan

Hasil dari pembinaan korpus dan analisis frekuensi perkataan dalam dua buku teks STAM adalah seperti berikut:

1. Kejayaan membina korpus ke atas dua buah buku teks di peringkat STAM sebagai projek awal memungkinkan pembinaan korpus yang lebih menyeluruh bagi mencakupi semua buku teks di peringkat tersebut yang berjumlah 19 buah semuanya bagi 12 subjek. Korpus yang dibina boleh diprogramkan agar boleh dimuat turun di makmal komputer di setiap sekolah menengah bagi pelajar mempelajari dan membuat kajian tentang penggunaan bahasa Arab di peringkat STAM bukan sahaja untuk mengetahui aspek kosa kata dan kata kunci tetapi juga aspek morfologi, sintaksis, gaya bahasa dan sebagainya.
2. Namun beberapa isu mungkin timbul semasa proses pembinaan korpus dari aspek pengkodan data kerana sifat bahasa Arab yang agak kompleks. Bagi membolehkan analisis mendalam dilakukan ke atas korpus yang dibina, proses pengkodan data haruslah melibatkan proses perletakan tatabahasa (*grammatical tagging*) bagi setiap perkataan yang terdapat dalam korpus agar perisian yang digunakan semasa menganalisis data dapat membezakan antara setiap perkataan terutama yang mempunyai bentuk fizikal yang sama. Adalah hampir mustahil untuk melakukan proses ini secara manual kerana sesuatu korpus mungkin mengandungi berjuta perkataan. Oleh itu pendekatan proses ini secara automatik harus diterokai dengan lebih mendalam.
3. Dapatan dari analisis korpus menunjukkan bahawa perkataan-perkataan yang kerap digunakan dan berada di senarai frekuensi yang tinggi adalah merujuk kepada jenis perbincangan di dalam buku-buku yang dikaji. Dalam

buku teks 1, perkataan seperti *قال* dan *نبي/رسول الله صلى الله عليه وسلم* di antara perkataan yang paling kerap digunakan. Situasi yang sama berlaku dalam buku teks 2. Perkataan-perkataan yang berada di senarai tertinggi merujuk kepada skop perbincangan tentang hukum-hakam seperti perkataan *دية*, *زوج*, *ولد* dan *يمين*. Ia seterusnya menjadi kosa kata penting dalam tajuk perkahwinan, wasiat dan jenayah di dalam buku tersebut.

4. Kewujudan perkataan yang merujuk kepada baris dalam buku teks 1 seperti *مهملة*, *كسر*, *فتح*, *ضم* dan *معجمة* menunjukkan bahawa ia amat mengambil berat tentang ketepatan sebutan beberapa perkataan penting kerana kesalahan sebutan akan membawa kepada kesalahan maksud. Ini berbeza dengan buku teks 2 yang tidak menekankan ciri ini. Sebaliknya buku teks 2 banyak merujuk kepada penulis *Matan Abu Shuja'* dengan kekerapan yang tinggi pada penggunaan perkataan *مصنف*. Selain daripada itu, pada setiap permulaan subtajuk, penulis buku ini banyak memulakannya dengan perkataan *تنبيه* yang untuk menarik perhatian pembaca.

5. Dari aspek kata nama, dapatan menunjukkan bahawa buku teks 1 amat mementingkan ketepatan sebutan sesuatu huruf seperti yang telah dijelaskan. Sementara situasi yang berbeza di dalam buku teks 2, di mana kata nama yang kerap digunakan sama ada lebih menjurus kepada tajuk yang dibincangkan atau bersifat umum. Dalam proses P&P subjek tersebut, perkataan yang menjurus kepada tajuk-tajuk tertentu mesti diberi keutamaan untuk difahami tanpa mengabaikan pemahaman perkataan umum.

6. Dalam aspek kata kerja, dapatan menunjukkan bahawa buku teks 1 didominasi oleh perkataan-perkataan yang membawa maksud 'berkata dan bercakap' dalam pelbagai variasi. Ini sesuai dengan sifat hadith yang disebarkan melalui percakapan sebelum penulisannya. Dalam buku teks 2, kata kerja-kata kerja yang kerap digunakan tidak merujuk kepada mana-mana tajuk yang dibincangkan kecuali beberapa perkataan. Sebaliknya, kata kerja yang merujuk kepada perkataan hukum dan umum kepada semua tajuk fiqh lebih banyak digunakan.

7. Analisis kata kunci bagi kedua-dua buku teks menunjukkan bahawa buku teks 1 banyak merujuk kepada ucapan hadith Rasulullah s.a.w kerana ia menjadi topik utama perbincangan. Sementara dalam buku teks 2, perkataan-perkataan yang menjurus kepada sesuatu tajuk fiqh yang dibincangkan menjadi kata kunci buku tersebut.

8. Proses P&P bagi dua subjek ini harus menekankan kepada pemahaman pelajar terhadap perkataan-perkataan penting yang kerap digunakan. Perkataan yang mesti diberi lebih keutamaan ialah perkataan yang menjurus kepada sesuatu tajuk walaupun kekerapan penggunaannya bukan pada kedudukan tertinggi. Selain daripada itu, perkataan-perkataan yang menjadi kata kunci sesuatu subjek juga harus diberi penekanan sewajarnya.

Rujukan

- Al-Sulaiti, L & Atwell, E. (2006). The Design of a Corpus of Contemporary Arabic. *International Journal of Corpus Linguistics*, 11 (2), 135-171.
- Al-Sulaiti, L. (3 Oktober 2013). Arabic corpora (Home page). Dicapai dari http://www.comp.leeds.ac.uk/eric/latifa/arabic_corpora.htm
- ELDA. (3 Oktober 2013). An-Nahar newspaper text corpus (Home page). <http://www.elda.org/catalogue/en/text/W0027.html>
- Fotos, J. T. (1931). Word and Idiom Frequency Counts in French and Their Value. *The Modern Language Journal*, 15 (5), 344-353.
- Hofland, K. & Jahansson, S. (1990). *Frequency Analysis of English Vocabulary and Grammar*. New York: Oxford University Press.
- Llantada, C. P. (2009). Textual, Genre and Social Features of Spoken Grammar. *Language Learning Technology*, 13 (1), 40-58.
- McEnergy, T & Xiao. R. (2011). What Corpora Can Offer in Language Teaching and Learning. Dalam Hinkel, *Handbook of Research in Second Language Teaching and Learning*. London: Routledge.
- Mustapha Ruslan. (1982). *al-Mufradat al-Sha'i'ah fi Kitabat Talamidh al-Saff al-Sadis al-Ibtidai*. Master thesis, Ain Shams University, Cairo.
- Na, L. & Nation, P. (1985). Factors Affecting Guessing Vocabulary in Context. *RELC Journal* 16 (1), 33-42.

- Nagy, W. E. and R. C. Anderson. (1984). How Many Words Are There in Printed School English?. *Reading Research Quarterly* 19 (3), 304-330.
- Nation, P. & Waring, R. (n.d.). Vocabulary Size, Text coverage and Word Lists. Dicapai dari <http://www.fltr.ucl.ac.be/fltr/germ/etan/bibs/vocab/cup.html>
- Nur Zafirah Mat Zuki. (2011). *Macrifah al-Ficl al-Mazid fi Surah al-Muzammil lada Talabah Kulliyah al-Quran wa al-Sunnah*. Projek Ilmiah, USIM, Nilai.
- Xue, G. & Nation P. (1984). A university word list. *Language Learning and Communication* 3 (2), 215-229.
- Wilkinson. M. (2011). WordSmith tools: The best corpus analysis program for translators?. *Translation journal*, 15 (3). Dicapai dari <http://translationjournal.net/journal/57corpus.htm>

Avoice of
Academia





A Voice of Academia

Academic Series of Universiti Teknologi MARA Cawangan Kedah

