



05-4506832



pustaka.upsi.edu.my



Perpustakaan Tuanku Bainun
Kampus Sultan Abdul Jalil Shah



PustakaTBainun



ptbupsi

HUBUNGAN ANALISIS KUANTITATIF ANTARA LEKSIKAL DENGAN METAFORA DALAM MANUSKRIP SENI BINA MELAYU



05-4506832

MUHAMAD FADZLLAH BIN HJ ZAINI



ptbupsi

UNIVERSITI PENDIDIKAN SULTAN IDRIS

2021



05-4506832



pustaka.upsi.edu.my



Perpustakaan Tuanku Bainun
Kampus Sultan Abdul Jalil Shah



PustakaTBainun



ptbupsi



**HUBUNGAN ANALISIS KUANTITATIF ANTARA LEKSIKAL DENGAN METAFORA
DALAM MANUSKRIP SENI BINA MELAYU**

MUHAMAD FADZLLAH BIN HJ ZAINI



**TESIS DIKEMUKAKAN BAGI MEMENUHI SYARAT UNTUK MEMPEROLEH IJAZAH
DOKTOR FALSAFAH**

**FAKULTI BAHASA DAN KOMUNIKASI
UNIVERSITI PENDIDIKAN SULTAN IDRIS**

2021





Sila tanda (✓)

Kertas Projek

Sarjana Penyelidikan

Sarjana Penyelidikan dan Kerja Kursus

Doktor Falsafah

✓

INSTITUT PENGAJIAN SISWAZAH

PERAKUAN KEASLIAN PENULISAN

Perakuan ini telah dibuat pada⁶(hari bulan).....⁷..... (bulan) 20²¹.....

i. Perakuan pelajar :

Saya, MUHAMAD FADZLLAH BIN HJ ZAINI (SILA NYATAKAN NAMA PELAJAR, NO. MATRIK DAN FAKULTI) dengan ini mengaku bahawa disertasi/tesis yang bertajuk HUBUNGAN ANALISIS KUANTITATIF ANTARA LEKSIKAL DENGAN METAFORA DALAM MANUSKRIP SENI BINA MELAYU

adalah hasil kerja saya sendiri. Saya tidak memplagiat dan apa-apa penggunaan mana-mana hasil kerja yang mengandungi hak cipta telah dilakukan secara urusan yang wajar dan bagi maksud yang dibenarkan dan apa-apa petikan, ekstrak, rujukan atau pengeluaran semula daripada atau kepada mana-mana hasil kerja yang mengandungi hak cipta telah dinyatakan dengan se jelasnya dan secukupnya

fadzllahzaini
Tandatangan pelajar

ii. Perakuan Penyelia:

Saya, PROF. MADYA DR. ANIDA SARUDIN (NAMA PENYELIA) dengan ini mengesahkan bahawa hasil kerja pelajar yang bertajuk HUBUNGAN ANALISIS KUANTITATIF ANTARA LEKSIKAL DENGAN METAFORA DALAM MANUSKRIP SENI BINA MELAYU

(TAJUK) dihasilkan oleh pelajar seperti nama di atas, dan telah diserahkan kepada Institut Pengajian Siswazah bagi memenuhi sebahagian/sepenuhnya syarat untuk memperoleh Ijazah DOKTOR FALSAFAH PENDIDIKAN BAHASA MELAYU (SLA NYATAKAN NAMA IJAZAH).

07/07/2021

Tarikh

Anida Sarudin
Tandatangan Penyelia
Prof. Madya Dr. Anida Sarudin
Jabatan Bahasa dan Kesusasteraan Melayu
Fakulti Bahasa dan Komunikasi
Universiti Pendidikan Sultan Idris
35900, Tg. Malim, Perak.



**INSTITUT PENGAJIAN SISWAZAH /
INSTITUTE OF GRADUATE STUDIES**

**BORANG PENGESAHAN PENYERAHAN TESIS/DISERTASI/LAPORAN KERTAS PROJEK
DECLARATION OF THESIS/DISSERTATION/PROJECT PAPER FORM**

Tajuk / Title: HUBUNGAN ANALISIS KUANTITATIF ANTARA LEKSIKAL DENGAN METAFORA
DALAM MANUSKRIP SENI BINA MELAYU

No. Matrik / *Matric's No.*: P20182002235

Saya / I: MUHAMAD FADZLLAH BIN HJ ZAINI

(Nama pelajar / *Student's Name*)

mengaku membenarkan Tesis/Disertasi/Laporan Kertas Projek (Kedoktoran/Sarjana)* ini disimpan di Universiti Pendidikan Sultan Idris (Perpustakaan Tuanku Bainun) dengan syarat-syarat kegunaan seperti berikut:-

acknowledged that Universiti Pendidikan Sultan Idris (Tuanku Bainun Library) reserves the right as follows:-

1. Tesis/Disertasi/Laporan Kertas Projek ini adalah hak milik UPSI.
The thesis is the property of Universiti Pendidikan Sultan Idris
2. Perpustakaan Tuanku Bainun dibenarkan membuat salinan untuk tujuan rujukan dan penyelidikan.
Tuanku Bainun Library has the right to make copies for the purpose of reference and research.
3. Perpustakaan dibenarkan membuat salinan Tesis/Disertasi ini sebagai bahan pertukaran antara Institusi Pengajian Tinggi.
The Library has the right to make copies of the thesis for academic exchange.

4. Sila tandakan (☒) bagi pilihan kategori di bawah / *Please tick (☒) for category below:-*

☐ **SULIT/CONFIDENTIAL**

Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub dalam Akta Rahsia Rasmi 1972. / *Contains confidential information under the Official Secret Act 1972*

☐ **TERHAD/RESTRICTED**

Mengandungi maklumat terhad yang telah ditentukan oleh organisasi/badan di mana penyelidikan ini dijalankan. / *Contains restricted information as specified by the organization where research was done.*

☒ **TIDAK TERHAD / OPEN ACCESS**

fadzllahzaini
(Tandatangan Pelajar/ Signature)

Tarikh: 07/07/2021

Anida Sarudin
(Tandatangan Penyele / *Signature of Supervisor*)
& (Nama & Cop Rasmi / *Name & Official Stamp*)
Prof. Madya Dr. Anida Sarudin
Jabatan Bahasa dan Kesusasteraan Melayu
Fakulti Bahasa dan Komunikasi
Universiti Pendidikan Sultan Idris
35900, Tg. Malim, Perak.

Catatan: Jika Tesis/Disertasi ini **SULIT @ TERHAD**, sila lampirkan surat daripada pihak berkuasa/organisasi berkenaan dengan menyatakan sekali sebab dan tempoh laporan ini perlu dikelaskan sebagai **SULIT** dan **TERHAD**.

Notes: If the thesis is CONFIDENTIAL or RESTRICTED, please attach with the letter from the organization with period and reasons for confidentiality or restriction.



PENGHARGAAN

Sekalung penghargaan dan ucapan terima kasih yang tidak terhingga ditujukan kepada Prof. Madya Dr. Anida Sarudin, selaku penyelia utama tesis ini. Beliau telah banyak memberi tunjuk ajar serta berkongsi ilmu dan bimbingan yang membina untuk menyelia penulisan ini. Segala bantuan dan serta bimbingan yang dihulurkan akan sentiasa menjadi kenangan dan terus menerus sebagai pembakar semangat diri saya. Di samping itu, saya tujukan khas penghargaan buat Prof. Madya Dr. Mazura @ Mastura Muhammad yang merupakan penyelia bersama yang memberi panduan dan galakan semasa kajian ini dijalankan sehingga selesai dengan cemerlang. Selain itu, Almarhumah Prof. Madya Dr. Siti Saniah Abu Bakar merupakan penyelia bersama yang sentiasa memberikan dorongan dalam penyelidikan dan segala bantuan kerohanian sehingga berjaya menghasilkan tesis ini. Buat ketiga-tiga penyelia saya, saya titipkan doa kejayaan di dunia dan akhirat. Semoga Allah S.W.T mencatat segala kebaikan penyelia.

Teristimewa buat Puan Hjh. Rosalmah Jamaludin Dan Tuan Hj. Zaini Hj. Jamari yang banyak memberi sokongan dan dorongan yang tinggi kepada anakanda kerana berjaya menyiapkan tesis ini. Perjalanan Ph.D ini atas doa restu ayahanda dan bonda yang tidak pernah putus merupakan pembakar semangat dan perlindungan terbaik buat anakanda, jutaan terima kasih anakanda ucapkan semoga Allah sentiasa memberi rahmat dan kebahagiaan buat kita sekeluarga. Terima kasih dengan tulus ikhlas ditujukan buat adik beradik, Mohd Fadzlullah, Nooradzahara, Mohd Fadzlillah & Suhaida, dan Nooradzaharah. Salam sayang buat Ammar Afeef dan Adeeb Azfar untuk berjaya pada masa hadapan.

Sekilas buat penghargaan bagi Kementerian Pendidikan Malaysia di atas manfaat Geran FRGS 2019-0036-108-02 kerana menjadi penaja utama dan tesis ini sebagai syarat wajib geran. Penghargaan ditujukan buat Prof. Madya Dr. Zukifli, Prof. Madya Dr. Anas, Dr. Husna Faredza dan Dr. Adenan. Seterusnya, penghargaan buat Prof. Madya. Dr. Mohd Rashid (Ketua Jabatan) Jabatan Bahasa dan Kesusasteraan Melayu, Fakulti Bahasa dan Komunikasi khususnya dan amnya Institut Pengajian Siswazah, Universiti Pendidikan Sultan Idris. Segala kebaikan dan ketulusan hati dalam bidang ilmu melahirkan manusia yang bijak seterusnya membentuk pengalaman yang terbaik. Ucapan terima kasih juga buat Flora, Nazrul, Amir, Lala, Izzue, Azim, Amer, Farhan dan rakan-rakan yang menjadi teman setia yang menyokong hingga berjaya. Sekalung penghargaan ditujukan buat Aswat, Effa, Parames, Dora, Nada, Athira, Grace dan Nabila (GRA FEB 2019/2020). Tulus ikhlas buat Prof. Dr. Nor Hashimah, Encik Tim, Cikgu Zabedah, dan Cikgu Ridzwan.





ABSTRAK

Kajian ini bertujuan untuk (1) mengenal pasti ketumpatan leksikal; (2) menganalisis hubungan matriks antara nod dengan leksikal; (3) menganalisis leksikogramatikal metafora dan kebarangkalian metafora dalam data korpus berdasarkan senarai kata kandungan terpilih; dan (4) menghuraikan dimensi variasi metafora berdasarkan pemboleh ubah topik dan sarana. Kajian ini menggunakan data Korpus Petua Membina Rumah (KPMR). Reka bentuk kajian ini ialah kajian kuantitatif, iaitu Pendekatan Statistik Korpus Linguistik. Kajian ini menggunakan kerangka teoretikal Statistik Korpus Linguistik dan Prosedur Pengenalpastian Metafora. Dapatan kajian ini dibahagikan kepada empat perkara, iaitu (1) ketumpatan leksikal menunjukkan fenomena kata yang berlaku melalui kekerapan tinggi, sederhana, rendah dan *hapaxes*; (2) representatif hubungan antara sepuluh nod dengan leksikal berlaku dan diplotkan dalam *GraphCollo* bagi memperoleh visualisasi geomatri ko-kejadian leksikal; (3) kehadiran nod kata kandungan menjelaskan pola leksikogramatikal, iaitu metafora yang terdapat dalam kolokasi dan *Key Word In Context* (KWIC). Nilai kebarangkalian metafora nod *tiang* [40%], *beroleh* [58%], *jahat* [14%] dan kolokasi bermotivasi untuk menjadi leksikal metafora (MUW) sebagai ekspresi metafora. Nod *beroleh* telah menunjukkan pertimbangan positif selaku leksikal metafora (MUW) berpotensi tinggi; (4) dimensi variasi leksikal yang digunakan untuk mengekspresikan leksikal metafora terdiri daripada dua metafora topik, empat metafora sarana. Kesimpulannya, statistik korpus linguistik dan prosedur pengenalpastian metafora yang digunakan dapat menzahirkan hubungan antara leksikal dengan metafora dalam manuskrip Seni Bina Melayu. Hubungan ini merujuk kepada bentuk *GraphCollo* yang diplotkan menjadi rajah geomatri nod dengan leksikal. Selain itu, kajian ini juga membentuk leksikal metafora (MUW) sebagai ekspresi metafora dalam manuskrip Seni Bina Melayu. Implikasi kajian membawa kepada signifikan perisian sebagai mesin boleh dibaca (*machine-readable*) bagi telaah teks klasik yang bertepatan sebagai revolusi integrasi penyelidikan pendidikan terkini.





QUANTITATIVE ANALYSIS OF RELATIONSHIP BETWEEN LEXICAL WITH METAPHOR IN MANUSCRIPT OF MALAY ARCHITECTURE

ABSTRACT

This study aims to (1) identify lexical density; (2) analyse the matrix relationship between nodes and lexical, 3) analyse metaphor lexicogrammatical and metaphor probability in corpus data based on selected content word lists; and (4) describe the dimensions of metaphorical variation based on topic variables and suggestions. This study uses the data from the House Building Tips (HBT) Corpus. The design of this quantitative study is based on the Linguistic Corpus Statistical Approach. This study uses the theoretical framework built based on the Linguistic Corpus Statistics and Metaphor Identification Procedures. The findings of this study are divided into four items namely (1) lexical density which indicates the word phenomena that occur through the high, medium, low and *hapaxes* frequencies; (2) relationship representation between ten nodes and lexical which occur and were plotted in *GraphCollo* to obtain a geometric visualisation of lexical co-occurrences; (3) the presence of word content's nodes which explain the lexicogrammatical patterns, that are metaphors found in collocations and Key Word in Context (KWIC). The metaphor probability value of the node 'pillar' (*tiang*) [40%], 'obtain' (*beroleh*) [58%], 'evil' (*jahat*) [14%] and collocation are stimulated to become metaphorical lexical (MUW) as metaphorical expressions. The node *beroleh* has shown positive consideration as a highly potential metaphorical lexical (MUW); (4) lexical variation dimensions utilised to express metaphorical lexical comprise of two topic metaphors and four suggested metaphors. To conclude, the statistical corpus linguistics and metaphor identification procedures used help to disclose the relationship between the lexical and metaphor in Malay Architecture manuscripts. This relationship refers to the *GraphCollo* design which was plotted into geometrical figure node with lexical. Also, this study shapes MUW as metaphorical expressions in Malay Architecture manuscripts. The study implications lead to the significance of software as a machine-readable for the study of classical texts which coincides with the recent educational revolution.



KANDUNGAN

Muka Surat

PENGAKUAN KEASLIAN PENULISAN	ii
PENGESAHAN PENYERAHAN TESIS	iii
PENGHARGAAN	iv
ABSTRAK	v
ABSTRACT	vi
KANDUNGAN	vii
SENARAI JADUAL	xiii
SENARAI RAJAH	xix
SENARAI SINGKATAN	xxv
SENARAI LAMPIRAN	xxvi

BAB 1 PENDAHULUAN

1.1	Pengenalan	1
1.2	Latar Belakang	2
1.3	Permasalahan Kajian	12
1.3.1	Kajian Teks Klasik (Manuskrip)	12
1.3.2	Kajian Metafora	15

1.3.3	Kajian Analisis Berpandukan Korpus	18
1.3.4	Fenomena di Malaysia Kajian Teks Klasik-Metafora- Analisis Berpandukan Korpus	20
1.4	Objektif Kajian	25
1.5	Soalan Kajian	27
1.6	Kepentingan Kajian	28
1.6.1	Kemajuan Ilmu dalam Teks Klasik	29
1.6.2	Kesejajaran Analisis Berpandukan Korpus Terhadap Prosedur Pengenalpastian Metafora	30
1.6.3	Nilai Tambah Literatur dalam Linguistik Bahasa Melayu	32
1.6.4	Transformasi Keterbacaan Manual Kepada Automatik	34
1.6.5	Representasi Ilmu Seni Bina : Sebagai Manfaat Sains dan Teknologi Bangunan	35
1.7	Definisi Operasional	37
1.7.1	Teks Klasik (Manuskrip)	37
1.7.2	Ketumpatan Leksikal	38
1.7.2.1	Analisis Berpandukan Korpus	39
1.7.2.2	Representasi Leksikal	41
1.7.3	Metafora	42
1.7.3.1	Prosedur Pengenalpastian Metafora	43
1.7.4	Ilmu Seni Bina Melayu	44
1.8	Kesimpulan	45

BAB 2 BIBLIOMETRIK ILMU LITERATUR

2.1	Pengenalan	46
-----	------------	----

2.2	Analisis Bibliometrik Kajian Literatur	47
2.3	Metodologi Analisis Bibliometrik	49
2.4	Dapatan Analisis Bibliometrik	50
2.4.1	<i>Neywords [text classic]</i>	50
2.4.2	<i>Neywords [metaphor]</i>	67
2.4.3	<i>Neywords [corpus driven]</i>	79
2.5	Analisis Kritis Merentas Trend Penyelidikan	88
2.5.1	Teks Klasik (Manuskrip)	88
2.5.2	Metafora	92
2.5.3	Analisis Berpandukan Korpus	96
2.6	Fenomena Kajian Manuskrip, Metafora dan Korpus Linguistik di Malaysia	100
2.7	Kerangka Teoretikal : Statistik Korpus Linguistik dan Metafora	107
2.8	Kesimpulan	112

BAB 3 METODOLOGI

3.1	Pengenalan	114
3.2	Deskripsi Data	115
3.2.1	Manuskrip Petua Membina Rumah	117
3.2.2	Korpus Rujukan	121
3.2.3	Korpus Petua Membina Rumah (KPMR)	124
3.3	Metodologi dan Pengumpulan Data	131
3.3.1	Reka Bentuk Penyelidikan	133
3.3.2	Kerangka Konseptual	148
3.4	Instrumen Kajian	151

3.4.1	Alat Statistik	152
3.4.2	LancsBox 4.5	154
3.4.3	Prosedur Pengenalpastian Metafora	162
3.4.3.1	Prosedur dan Penjelasan	163
3.4.3.2	Leksikogramatikal Metafora	165
3.4.3.3	Kebarangkalian Metafora	167
3.4.3.4	Dimensi Variasi Metafora	169
3.5	Analisis Data	171
3.5.1	Penyelarasan Nilai Ambang	175
3.6	Kesahan dan Kebolehpercayaan	177
3.6.1	Kesahan	178
3.6.2	Kebolehpercayaan	179
3.7	Batasan Kajian	182
3.8	Kesimpulan	184

BAB 4 TABURAN KATA

4.1	Pengenalan	185
4.2	Kekerapan Kata KPMR	186
4.2.1	<i>Hapaxes</i>	214
4.3	Senarai Kata Kunci	226
4.4	Senarai Kata Kandungan	254
4.5	Kesimpulan	271

BAB 5 LEKSIKOGRAMATIKAL METAFORA

5.1	Pengenalan	272
-----	------------	-----

5.2	Nod <i>tiang</i>	275
5.3	Nod <i>beroleh</i>	318
5.4	Nod <i>jahat</i>	365
5.5	Pendidikan Bermula daripada Awal Pembinaan Rumah	394
5.5	Kesimpulan	419

BAB 6 DIMENSI VARIASI METAFORA

6.1	Pengenalan	421
6.2	Kebarangkalian Metafora <i>tiang</i>	423
6.3	Kebarangkalian Metafora <i>beroleh</i>	434
6.4	Kebarangkalian Metafora <i>jahat</i>	445
6.5	Dimensi Variasi Metafora	453
6.5.1	Metafora Topik : Manusia	453
6.5.2	Metafora Topik : Abstrak	459
6.5.3	Metafora Sarana : Pergerakan/ Kedudukan	467
6.5.4	Metafora Sarana : Objek/Bangunan	470
6.5.5	Metafora Sarana : lain-lain	475
6.5.6	Kata Sarana : Kata Kerja	479
6.6	Kesimpulan	482

BAB 7 KESIMPULAN DAN CADANGAN

7.1	Pengenalan	483
7.2	Penilaian Objektif Kajian	484
7.2.1	Mengenal pasti kekerapan dan taburan kata dalam data KPMR	485

7.2.2	Menganalisis penghubungan matriks antara leksikal dalam data KPMR	487
7.2.3	Menganalisis leksikogramartikal metafora dalam data KPMR	489
7.2.4	Menghuraikan dimensi variasi metafora berdasarkan pemboleh ubah metafora	491
7.3	Rumusan Dapatan Kajian	493
7.4	Sumbangan Kajian	496
7.4.1	Ketumpatan leksikal perspektif teks klasik atau manuskrip Melayu	497
7.4.2	Literatur ilmu Melayu terdahulu, kini dan masa depan	497
7.4.3	Mesin keterbacaan representasi ilmu teks klasik atau manuskrip Melayu	498
7.4.4	Representasi ilmu seni bina	499
7.4.5	Pendidikan Bahasa Melayu Sektor TVET	501
7.5	Saranan Masa Muka	502
7.5.1	Kajian statistik leksikal manuskrip terkini	502
7.5.2	Kajian manuskrip bukan sastera sebagai khazanah ilmu	503
7.5.3	Kajian variasi dialek sebagai pengumpulan perbendaharaan leksikal	504
7.6	Kesimpulan	504
	RUJUKAN	506
	LAMPIRAN	529



SENARAI JADUAL

No. Jadual		Muka Surat
2.1	Maklumat umum dalam Springer Link	49
2.2	Jenis Dokumen <i>keyword [text classic]</i>	51
2.3	Jenis Disiplin <i>keyword [text classic]</i>	52
2.4	Senarai Subdisiplin <i>keyword [text classic]</i>	56
2.5	Senarai Dokumen <i>keyword [metaphor]</i>	67
2.6	Senarai Disiplin <i>keywords [metaphor]</i>	68
2.7	Senarai Subdisiplin <i>keywords [metaphor]</i>	72
2.8	Senarai dokumen <i>keywords [corpus driven]</i>	80
2.9	Senarai disiplin <i>keyword [corpus driven]</i>	81
3.1	Maklumat korpus rujukan Korpus UKM-DBP	122
3.2	Maklumat teknikal manuskrip	128
3.3	Statistik Asas KPMR	129
3.4	Rumusan tatacara pengumpulan data	132
3.5	Kekerapan Pemerhatian (<i>Observed Frequencies</i>)	142
3.6	Rumusan daripada pengumpulan data	173
3.7	Kesahan dan kebolehpercayaan daripada aspek kajian ini	181
4.1	Rumusan senarai leksikal berbeza berdasarkan nilai relatif dan mentah	199



4.2	Representasi <i>hapaxes</i> dalam KWIC	220
4.3	KWIC variasi leksikal <i>ukur</i>	222
4.4	KWIC variasi leksikal <i>jauh</i>	223
4.5	KWIC variasi leksikal <i>belah</i>	224
4.6	KWIC variasi leksikal <i>kelahi</i>	224
4.7	KWIC variasi leksikal <i>bayang</i>	225
4.8	KWIC variasi leksikal <i>likur</i>	225
4.9	Representasi 100 kata kunci	227
4.10	Representasi kata kunci positif dan negatif	236
4.11	Kolokasi leksikal kata kunci ' <i>beroleh</i> '	240
4.12	Leksikal <i>mendirikan</i> + kolokasi	246
4.13	Leksikal <i>empunya</i> + kolokasi	248
4.14	Leksikal <i>alamat</i> + kolokasi	250
4.15	Leksikal <i>bab</i> + kolokasi	252
4.16	Representasi senarai kata kandungan lingkungan 100 leksikal teratas	254
4.17	Representasi nod <i>tiang</i> dalam KWIC	258
4.18	Representasi nod <i>ukuran</i> dalam KWIC	259
4.19	Representasi nod <i>rumah</i> dalam KWIC	259
4.20	Representasi nod <i>lubang</i> dalam KWIC	260
4.21	Representasi nod <i>bendul</i> dalam KWIC	261
4.22	Representasi nod <i>beroleh</i> dalam KWIC	262
4.23	Representasi nod <i>mendirikan</i> dalam KWIC	263
4.24	Representasi nod <i>lipat</i> dalam KWIC	264
4.25	Representasi nod <i>bubuh</i> dalam KWIC	264



4.26	Representasi nod <i>perkelahian</i> dalam KWIC	265
4.27	Representasi nod <i>jahat</i> dalam KWIC	266
4.28	Representasi nod <i>senget</i> dalam KWIC	267
4.29	Representasi nod <i>dukacita</i> dalam KWIC	268
4.30	Representasi nod <i>baik</i> dalam KWIC	269
4.31	Representasi nod <i>celaka</i> dalam KWIC	270
5.1	Senarai 15 kata kandungan	274
5.2	Makna leksikal <i>tiang</i> berdasarkan Kamus Dewan Edisi Keempat	275
5.3	Representasi <i>hapaxes</i> kolokasi nod <i>tiang</i>	278
5.4	Representasi <i>hapaxes</i> kolokasi nod <i>tiang</i> (bukan <i>hapaxes</i> KPMR)	282
5.5	Representasi kolokasi nod <i>tiang</i> > 5.0	286
5.6	Representasi kolokasi nod <i>tiang</i> pengujian t-skor	292
5.7	Representasi nod <i>tiang</i> dalam KWIC ko-kejadian <i>tiang-baik-rumah</i>	296
5.8	Representasi nod <i>tiang</i> dalam KWIC ko-kejadian <i>tiang-ukuran-rumah</i>	297
5.9	Representasi kolokasi nod <i>tiang</i> pengujian mi-skor	300
5.10	Representasi <i>tiang-bertuah</i> dalam KWIC	304
5.11	Representasi <i>tiang-serambi</i> dalam KWIC	306
5.12	Representasi kolokasi nod <i>tiang</i> pengujian logDice	310
5.13	Representasi <i>tiang-lubang</i> dalam KWIC	314
5.14	Representasi <i>tiang-kayu</i> dalam KWIC	317
5.15	Representasi kolokasi nod <i>beroleh</i> dengan nilai <i>hapaxes</i>	320
5.16	Representasi <i>hapaxes</i> korpus keseluruhan dalam kolokasi nod <i>beroleh</i>	324





5.17	Representasi kolokasi nod <i>beroleh</i> > 5.0 (kekerapan mentah)	327
5.18	Representasi <i>beroleh-padahnya</i> dalam KWIC	331
5.19	Representasi kolokasi nod <i>beroleh</i> pengujian t-skor	334
5.20	Representasi <i>beroleh-pintu</i> dalam KWIC	340
5.21	Representasi <i>beroleh-harta</i> dalam KWIC	342
5.22	Representasi kolokasi leksikal nod <i>beroleh</i> pengujian mi-skor	345
5.23	Representasi <i>beroleh-kemuliaan</i> dalam KWIC	351
5.24	Representasi <i>beroleh-laba</i> dalam KWIC	353
5.25	Representasi senarai kolokasi nod <i>beroleh</i> pengujian logDice	356
5.26	Representasi <i>beroleh-hamba</i> dalam KWIC	361
5.27	Representasi <i>beroleh-anak</i> dalam KWIC	363
5.28	Representasi <i>hapaxes</i> KPMR kolokasi nod <i>jahat</i>	366
5.29	Representasi <i>hapaxes</i> kolokasi nod <i>jahat</i>	367
5.30	Representasi kolokasi nod <i>jahat</i> kekerapan mentah > 5.0	371
5.31	Representasi <i>jahat-bumi</i> dalam KWIC	374
5.32	Representasi kolokasi nod <i>jahat</i> pengujian t-skor	377
5.33	Representasi <i>jahat-tiada</i> dalam KWIC	381
5.34	Representasi kolokasi leksikal nod <i>jahat</i> pengujian mi-skor	383
5.35	Representasi <i>jahat-perkelahian</i> dalam KWIC	386
5.36	Representasi kolokasi leksikal nod <i>jahat</i> (logDice)	389
5.37	Representasi <i>jahat-diam</i> dalam KWIC	393





5.38	KWIC bagi nod <i>tiang</i> bersama nilai pendidikan	400
5.39	KWIC bagi nod <i>beroleh</i> bersama nilai pendidikan	406
5.40	KWIC bagi nod <i>jahat</i> bersama nilai pendidikan	409
5.41	KWIC bagi nod mendirikan bersama nilai pendidikan	413
5.42	KWIC bagi nod ukuran bersama nilai pendidikan	417
6.1	Leksikal yang representasi tinggi signifikan sebagai MUW nod <i>tiang</i>	424
6.2	Leksikal yang representasi sederhana signifikan sebagai MUW nod <i>tiang</i>	426
6.3	Leksikal yang representasi rendah signifikan sebagai MUW nod <i>tiang</i>	430
6.4	Leksikal yang representasi tinggi signifikan sebagai MUW nod <i>beroleh</i>	435
6.5	Leksikal yang representasi sederhana signifikan sebagai MUW nod <i>beroleh</i>	438
6.6	Leksikal yang representasi rendah signifikan sebagai MUW nod <i>beroleh</i>	442
6.7	Leksikal yang representasi tinggi signifikan sebagai MUW nod <i>jahat</i>	446
6.8	Leksikal yang representasi sederhana signifikan sebagai MUW nod <i>jahat</i>	449
6.9	Leksikal yang representasi rendah signifikan sebagai MUW nod <i>jahat</i>	450
6.10	Metafora Topik- Manusia dalam KWIC	458
6.11	Metafora Topik- Abstrak dalam KWIC	464
6.12	MS-Pergerakan/Kedudukan dalam KWIC	467
6.13	MS-Objek atau Bangunan dalam KWIC	471
6.14	MS- Lain-lain dalam KWIC	477





05-4506832



pustaka.upsi.edu.my



Perpustakaan Tuanku Bainun
Kampus Sultan Abdul Jalil Shah



PustakaTBainun



ptbupsi

xviii

6.15

Kata sarana : kata kerja dalam KWIC

480



05-4506832



pustaka.upsi.edu.my



Perpustakaan Tuanku Bainun
Kampus Sultan Abdul Jalil Shah



PustakaTBainun



ptbupsi



05-4506832



pustaka.upsi.edu.my



Perpustakaan Tuanku Bainun
Kampus Sultan Abdul Jalil Shah



PustakaTBainun



ptbupsi



SENARAI RAJAH

No. Rajah		Muka Surat
2.1	Carian <i>Neyword</i> [Text Classic]	55
2.2	Awan Kata <i>Neyword</i> [Text Classic]	59
2.3	Senarai Keperluan daripada Kelompangan Kajian	66
2.4	Awan Kata <i>Neyword</i> [Metaphor]	75
2.5	Pola Penyelidikan Topik Metafora dan Korpus	77
2.6	Awan Kata <i>Neyword</i> [Corpus Driven]	84
2.7	Pola Penyelidikan Bahasa dalam Topik Berpandukan Korpus	85
2.8	Kerangka Teoretikal Statistik Berpandukan Korpus dan Prosedur Pengenalpastian Metafora	110
3.1	Keratan Manuskrip Petua Membina Rumah MSS741	118
3.2	Elemen yang digugurkan dalam Proses Pembersihan	120
3.3	Elemen Doa dan Wafak dibersihkan	121
3.4	Contoh Transliterasi Rumi dalam simpanan <i>Plain Text</i> (.txt)	124
3.5	Data <i>engine.malaynlp</i>	125
3.6	Sampel Data Kitab Abu Ma'syar Al-Faliki	126
3.7	Sampel Data Kitab Tajul Muluk	126
3.8	Sampel Data MSS741	127





3.9	Sampel Data MSS1415	127
3.10	Sampel Data MSS1521	127
3.11	Sampel Data MSS1849	128
3.12	Sampel Data MSS2001	128
3.13	Tatacara Pengumpulan Data	131
3.14	Senarai Kekerapan Kata (Kekerapan Mentah)	135
3.15	Senarai Kekerapan Kata (Kekerapan Relatif)	136
3.16	Senarai Kata Kunci	138
3.17	Representasi Penghubungan Matriks Visualisasi Leksikal <i>rumah</i> (mi-Skor)	143
3.18	Representasi Penghubungan Matriks Visualisasi Leksikal <i>Rumah-Kebakaran-Itu</i> (mi-Skor)	144
3.19	Representasi Penghubungan Matriks Visualisasi Leksikal <i>Rumah</i> (t-Skor)	145
3.20	Representasi Penghubungan Matriks Visualisasi Leksikal <i>Rumah-Bulan-Mendirikan</i> (t-Skor)	145
3.21	Infografik Konseptual Aliran Kajian	149
3.22	Perisian Lancsbox 4.5	155
3.23	<i>KWIC Tool</i> Bagi Carian Leksikal <i>Rumah</i>	156
3.24	<i>Whelk Tool</i> Bagi Carian Leksikal <i>Rumah</i>	157
3.25	<i>Graphcoll Tool</i> Senarai Kolokasi Leksikal <i>Rumah-Bulan</i>	158
3.26	Visualisasi Kolokasi <i>Graphcoll Tool</i>	158
3.27	<i>Words Tool</i>	159
3.28	<i>N-Gram Tool</i>	160
3.29	<i>Text Tool</i>	161
3.30	Konsep Metafora	170





3.31	Nilai ambang t-skor	176
3.32	Nilai ambang mi-skor	176
3.33	Nilai ambang logDice	177
3.34	Nilai ambang <i>Simple Maths Procedure</i>	177
3.31	Infografik Batasan Kajian	182
4.1	Representasi trend kekerapan mentah KPMR	187
4.2	Representasi trend kekerapan relatif KPMR	188
4.3	Representasi kekerapan relatif 100 kali ke atas	190
4.4	Representasi kekerapan mentah 100 kali ke atas	191
4.5	Representasi kekerapan relatif 100 hingga 50 kali	192
4.6	Representasi kekerapan mentah 100 hingga 50 kali	193
4.7	Representasi kekerapan relatif 50 hingga 10 kali	195
4.8	Representasi kekerapan mentah 50 hingga 10 kali	196
4.9	Representasi kekerapan relatif leksikal 2 hingga 9 kali	198
4.10	Representasi kekerapan mentah leksikal 2 hingga 9 kali	198
4.11	Kedudukan representasi <i>hapaxes</i> (KM)	215
4.12	Kedudukan representasi <i>hapaxes</i> (KR)	215
4.13	Representasi kekerapan mentah <i>hapaxes</i>	216
4.14	Representasi kekerapan relatif <i>hapaxes</i>	217
4.15	Pecahan jenis <i>hapaxes</i>	217
4.16	Representasi signifikan leksikal berdasarkan <i>Simple Math Procedure</i> (SMP)	226
4.17	Representasi 100 leksikal yang berbeza	233
4.18	Representasi pertama 20 kata kunci	234
4.19	Representasi kolokasi geomatriks kata kunci (+)	238





4.20	Regresion Square (R^2) logDice kolokasi <i>beroleh</i>	243
4.21	Representasi kolokasi <i>beroleh</i> dengan lima kata kunci	245
4.22	Representasi kolokasi <i>mendirikan</i> dengan enam kata kunci	247
4.23	Representasi kolokasi <i>empunya</i> dengan tiga kata kunci	249
4.24	Representasi kolokasi <i>alamat</i> dengan lima kata kunci	251
4.25	Representasi kolokasi <i>bab</i> dengan empat kata kunci	253
5.1	Representasi kekerapan mentah kolokasi nod <i>tiang</i>	277
5.2	Representasi visual graf kolokasi nod <i>tiang</i> (kekerapan kolokasi)	288
5.3	Representasi visual graf kolokasi nod <i>tiang</i> (kekerapan kolokasi) plot <i>positional</i>	290
5.4	Representasi t-skor leksikal kolokasi nod <i>tiang</i>	291
5.5	Representasi t-skor leksikal berkolokasi nod <i>tiang</i> (> 3.0)	292
5.6	Representasi visual graf kolokasi nod <i>tiang</i> (t-skor)	295
5.7	Representasi graf kolokasi nod <i>tiang-rumah</i>	296
5.8	Representasi mi-skor leksikal kolokasi nod <i>tiang</i>	298
5.9	Representasi mi-skor leksikal kolokasi nod <i>tiang</i>	299
5.10	Representasi visual graf kolokasi nod <i>tiang</i> (mi-skor)	302
5.11	Representasi visual graf kolokasi nod <i>tiang</i> (mi-skor) L-R	303
5.12	Representasi visual graf kolokasi <i>tiang-bertuah</i>	304
5.13	Representasi visual graf kolokasi <i>tiang-serambi</i>	306
5.14	Representasi logDice bagi leksikal berkolokasi nod <i>tiang</i>	309
5.15	Representasi visual graf kolokasi nod <i>tiang</i> (logDice)	312





5.16	Representasi visual graf kolokasi nod <i>tiang</i> (logDice) L-R	312
5.17	Representasi <i>tiang-lubang</i> pengujian logDice	313
5.18	Representasi <i>tiang-kayu</i> pengujian logDice	317
5.19	Representasi kolokasi nod <i>beroleh</i> (kekerapan mentah)	319
5.20	Representasi graf visual kolokasi nod <i>beroleh</i> (kekerapan mentah)	329
5.21	Representasi graf visual kolokasi nod <i>beroleh</i> (kekerapan mentah) L-R	330
5.22	Representasi kolokasi nod <i>beroleh</i> pengujian t-skor	334
5.23	Representasi graf visualisasi kolokasi nod <i>beroleh</i> (t-skor)	337
5.24	Representasi graf visualisasi kolokasi nod <i>beroleh</i> (t-skor) L-R	338
5.25	Representasi graf visualisasi leksikal <i>beroleh-pintu</i>	339
5.26	Representasi graf visualisasi leksikal <i>beroleh-harta</i>	341
5.27	Representasi kolokasi nod <i>beroleh</i> dengan pengujian mi-skor	344
5.28	Representasi kolokasi nod <i>beroleh</i> (mi-skor)	348
5.29	Representasi kolokasi nod <i>beroleh</i> (mi-skor) L-R	349
5.30	Representasi <i>beroleh-kemuliaan</i>	350
5.31	Representasi <i>beroleh-laba</i>	352
5.32	Representasi kolokasi nod <i>beroleh</i> pengujian logDice	355
5.33	Representasi graf visualisasi kolokasi nod <i>beroleh</i> (logDice)	359
5.34	Representasi graf visualisasi kolokasi nod <i>beroleh</i> (logDice) L-R	359
5.35	Representasi <i>beroleh-hamba</i>	360



5.36	Representasi <i>beroleh-anak</i>	362
5.37	Representasi kolokasi leksikal nod <i>jahat</i>	365
5.38	Representasi graf visualisasi kolokasi nod <i>jahat</i> (kekerapan mentah)	372
5.39	Representasi graf visualisasi kolokasi nod <i>jahat</i> (kekerapan mentah) L-R	373
5.40	Representasi <i>jahat-bumi</i>	374
5.41	Representasi kolokasi nod <i>jahat</i> (t-skor)	376
5.42	Representasi graf visualisasi kolokasi nod <i>jahat</i> (t-skor)	378
5.43	Representasi graf visualisasi kolokasi nod <i>jahat</i> (t-skor) L-R	379
5.44	Representasi <i>jahat-tiada</i>	380
5.45	Representasi kolokasi leksikal nod <i>jahat</i> (mi-skor)	382
5.46	Representasi graf visualisasi kolokasi leksikal nod <i>jahat</i> (mi-skor)	384
5.47	Representasi graf visualisasi kolokasi leksikal nod <i>jahat</i> (mi-skor) L-R	385
5.48	Representasi <i>jahat-perkelahian</i>	386
5.49	Representasi kolokasi leksikal nod <i>jahat</i> pengujian logDice	388
5.50	Representasi graf visualisasi kolokasi leksikal nod <i>jahat</i> (logDice)	390
5.51	Representasi graf visualisasi kolokasi leksikal nod <i>jahat</i> (logDice) L-R	391
5.52	Representasi <i>jahat-diam</i>	392

SENARAI SINGKATAN

KBH	Korpus Berita Harian
KM	Kekerapan Mentah
KPMR	Korpus Petua Membina Rumah
KR	Kekerapan Relatif
L	<i>Left/Kiri</i>
MIP	<i>Metaphor Identification Procedure</i>
MUW	<i>Metaphorically Used Words/</i> Leksikal Metafora
R	<i>Right/Kanan</i>
SMP	<i>Simple Maths Procedure</i>
STAT	Statistik



SENARAI LAMPIRAN

- A Senarai Kata
- B Senarai Kata Kunci
- C Leksikogramatikal Metafora



BAB 1

PENDAHULUAN

1.1 Pengenalan

Tradisi penyelidikan dalam dunia akademik saban tahun telah memberikan satu impak yang baharu dalam usaha mencetuskan penemuan baharu dalam ruang lingkup penyelidikan yang diteliti . Sumbangan penyelidikan tersebut mewujudkan impak yang positif yang bertindak sebagai pemangkin bagi penyelesaian alternatif. Oleh hal demikian, bab ini berupaya mewujudkan ruang perbincangan ilmu secara ilmiah bagi memperoleh cerapan senario dunia penyelidikan dan inovasi dalam pendidikan. Bab ini diberikan kepada beberapa bahagian iaitu latar belakang, permasalahan kajian, objektif kajian, soalan kajian, kepentingan kajian, dan definisi operasional. Justeru, kupasan ini

menjadi mauduk kajian hubungan antara leksikal dan metafora dalam manuskrip Melayu seni bina.

1.2 Latar Belakang Kajian

Penyelidikan manuskrip merupakan satu cabang ilmu dalam meneroka perluasan ilmu yang terkandung dalam sesebuah teks klasik. Dalam konteks penyelidikan dunia, penerokaan data ini masih dikaji dan digunakan sebagai bahan utama bagi penyelesaian serta rujukan. Teks klasik bukan sahaja menjadi nadi kehebatan sesebuah tamadun sebaliknya memerihalkan soal sains dan teknologi yang digunakan semasa kegemilangan negara bangsa (Ricciardi, 2019). Daripada tahun 2014 sehingga 2018 *Springer Link* telah menerbitkan buku bersiri dalam kajian teks klasik dalam sains yang mengumpulkan kajian sejarah, perkembangan sains, pengaruh budaya seterusnya latar belakang aspek ketamadunan dalam sains. Hubungan antara teks ini dalam membentuk satu manuskrip menjadi dasar kepada sesebuah naratif yang mengandungi makna implisit (Strowe, 2019). Di samping itu, naratif ini disampaikan mengikut bahasa tamadun yang menjadi penulis perihal konteks dan ilmunya (Birch, 2019; Chen Bo, 2018; Deps & Charlier, 2019; Shu, 2018). Oleh itu, sumber bahasa menjadi rangkaian utama dalam memahami naratif yang disampaikan bagi sesebuah manuskrip bangsa.

Dalam konteks Malaysia, iaitu masyarakat Melayu feudal yang tidak banyak saluran maklumat, manuskrip Melayu merupakan sumber peribadi bertulis yang rakyat



warisi sebelum dan setelah kedatangan orang Eropah ke dunia Melayu (Ding Choo Ming, 2016). Kepelbagaian catatan yang diusahakan mendorong lahirnya genre manuskrip Melayu seperti naskhah dan karya sastera klasik (Yakob, 2018). Isi kandung manuskrip Melayu sangat kaya dengan meliputi ilmu bintang, penujuman, ubat, tadbir mimpi, kias, raksi, berahi, bahasa, pertukangan, pertanian, perhitungan, undang-undang, pelayaran, agama dan pelbagai lagi (Ding Choo Ming, 2016; Ming, Chambert-Loir, & Pudjiastuti, 2014; Yakob, 2018). Dalam pada itu, naskhah Melayu merupakan satu dokumen sosial yang harus diketengahkan ke persada masyarakat agar sejarah dan rakaman peristiwa masa silam kekal menjadi bahan ingatan dan pengajaran selamanya (Jelani Harun & Rogayah, 2019). Sehubungan dengan itu, perspektif manuskrip Melayu perlu diteliti sebagai sumber maklumat umum kerana ruang perbincangan luas dan melangkau pelbagai disiplin ilmu.



Signifikan teks klasik atau manuskrip Melayu ini kepada pembangunan negara terutama pendidikan adalah tradisi penulisan agung yang tidak terputus sejak abad ke-13 hingga abad ke-19 di kebanyakan istana Melayu. Daripada itu, pemikir, pengarang, ulama, penasihat menjadi pakar rujuk raja-raja Melayu (Ding Choo Ming, 2016). Sumber rujukan ini menjadi medium kepada laman carian ilmu yang bersifat manual dan berkedudukan. Masyarakat sering didedahkan dengan kebudayaan Melayu yang telah difahami dalam konteks sejarah yang panjang serta ditandai oleh hubungan antarabangsa bersama negara asing. Tetapi, andaian ini meletakkan pengajian Dunia Melayu terdedah kepada tekanan pengaruh luar. Seolah-olah dunia Melayu tiada jiwa, roh dan jati diri yang berteraskan bahasa, sastera, budaya, masyarakat dan negara bangsa (Hawa, 1998). Andaian ini disangkal oleh sarjana falsafah Melayu. Hal ini



kerana sebelum manuskrip yang lahir daripada tradisi penulisan Melayu, sudah ada tradisi lisan Melayu (Ding Choo Ming, 2016; Hamzah, 2019; Hassan, 2016; Yakob, 2018).

Tradisi lisan Melayu yang lahir daripada kognitif penutur Melayu seperti ungkapan, pantun, lipur lara dan banyak lagi. Bahasa sangat dikaitkan dengan kognisi manusia dalam proses pemahaman dan penyampaian maklumat. Hal ini merupakan satu gambaran budaya yang seiringan dengan lahirnya bahasa. Maklumat manuskrip ini lahir daripada pemikiran pengarang yang tidak dapat memisahkan idealisme dan realisme (Hassan, 2016). Perkara ini merupakan satu naratif perjalanan yang panjang dan menjadi pedoman kepada masyarakat. Apabila menyatakan soal idealisme dan realisme, satu perkara yang tidak dapat dinafikan sikap masyarakat Melayu dahulunya adalah berjaya menggabungkan alam dan nilai teladan dalam kehidupan masyarakat.

Dalam Hikayat Merong Mahawangsa telah diceritakan kisah kehadiran Islam ke Dunia Melayu dengan cara yang rumit, namun menarik. Hal ini memberikan pandangan sumber maklumat yang menarik soal kewujudan Hinduisme dan Islam, satu contoh kefahaman pelbagai budaya dengan meliputi gejala yang baru wujud bersama-sama dengan yang lama (Hawa, 1998). Selain itu, Bustan al-Salatin merupakan satu naratif yang dikelaskan sebagai teks ketatanegaraan dan memperoleh cerminan ciri keislaman (Jelani, 2004). Naratif seumpama ini merupakan bingkisan umum kepengarangan serta tatacara penulis Melayu dahulu mengekalkan seterusnya menyampaikan maklumat yang penting sebagai kemandirian (survival). Di samping itu, Hikayat Hang Tuah yang dihasilkan pada abad ke-17 yang memiliki kandungan masih



relevan untuk dibincangkan dan diberikan perhatian (Kassim, 2008). Karya agung ini mempunyai nilai signifikan yang tersendiri dan masih dikaji sehingga hari ini .

Bagaimana manuskrip Melayu dalam bidang-bidang lain? Berdasarkan maklumat daripada Pusat Kebangsaan Manuskrip Melayu, Perpustakaan Negara Malaysia (PNM), pelbagai genre manuskrip yang disimpan sebagai khazanah negara. Antaranya kitab perubatan Melayu (Kitab Tib), Takbir Gempa, Rasi, Ilmu-ilmu Agama, kitab pertukangan (Petua Membina Rumah) dan lain-lain. Manuskrip sedemikian dipanggil sebagai bukan sastera atau bukan sastera kerana pemilikan ilmu dan tatacara penulisan sebagai pengalaman rakyat (Ding Choo Ming, 2016). Daripada hasil penulisan berbentuk manuskrip, khasnya naskhah-naskhah yang tergolong sebagai contoh ‘Kitab Tib’ dapat sedikit maklumat yang menunjukkan tradisi itu telah lama dikenali dan diamalkan dalam kehidupan masyarakat Melayu (Harun, 2006). Setelah dibuat kajian khusus berkaitan kepelbagaian isi dan mesej dalam manuskrip Melayu dari masa ke masa dapat mengikuti kedalaman dan keluasan ilmu masyarakat Melayu dari pelbagai bidang. Manuskrip Melayu merupakan satu-satunya data bank bertulis tentang akal budi, citra, rasa, dan kebijaksanaan orang Melayu sejak masa silam (Ding Choo Ming, 2016; Hassan, 2016; Nor Hashimah, 2013; Yakob, 2018).

Metafora merupakan satu aspek yang sering diteliti dalam kajian-kajian teks klasik atau manuskrip. Menurut Kamberi (2013), analisis metafora telah digunakan dalam kesusasteraan, terutamanya dalam puisi sebelum ini. Walau bagaimanapun, ia telah digunakan dalam pemerolehan bahasa kedua juga. Namun, Kamberi menjelaskan keberkesanan pengajian ini mampu memberi nilai pemikiran aras tinggi dalam





kemahiran berbahasa. Analisis metafora ini bukan sahaja melihat konteks makna yang disampaikan secara eksplisit dan implisit, tetapi melihat konteks leksikal yang digunakan bagi membentuk satu garisan metafora dalam semantik sesebuah wacana. Perspektif Aristotle telah menetapkan tiga tujahan dalam fitur metafora iaitu; (1) metafora berlaku pada peringkat leksikal, tidak pada peringkat ayat; (2) metafora sebagai penyelewengan daripada bahasa harfiah, Cruse (2000:202) menyebut sebagai *“the substitution view of metaphor”*; (3) metafora dibentuk berdasarkan ciri-ciri persamaan antara dua benda. Melalui metafora manusia cuba melihat sesuatu berdasarkan sesuatu yang lain (Hassan, 2016).

Definisi metafora ini merupakan sesuatu yang umum menerangkan fenomena manusia berkomunikasi dan berpotensi, berfikir tentang sesuatu dari segi sesuatu yang lain. Oleh hal demikian, Prosedur Pengenalpastian Metafora atau *Metaphor Identification Prosedure (MIP)* dibentuk bagi melihat proses pola leksikal yang digunakan bagi membentuk metafora (Pragglejaz, 2007). Semino (2008), merupakan salah seorang pengkaji dalam kumpulan Pragglejaz menyatakan wujud perbezaan *metaphor expressions* (leksikal kata nama ‘battle’ dalam ‘a battle of metaphor’) dan metafora konseptual (*ARGUMENT IS WAR*). *Conceptual Metaphor Theory (CMT)* yang dahulu dilihat sebagai realisasi linguistik yang terakhir. Sekarang adalah masa untuk menjadi lebih tepat mengenai apa yang dimaksudkan apabila membincangkan tentang ‘*linguistic metaphor*’ (metafora linguistik), atau ‘*metaphorical expressions*’ (ekspresi metafora) atau ‘*metaphorically used word (MUW)*’ (kata-kata yang digunakan metafora) (Pragglejaz, 2007; Semino, 2008). Hal ini adalah satu isu metafora cenderung untuk meninggalkan makna agak kabur, tetapi boleh ditangani secara lebih





jelas dan sistematik (Goatly, 1997; Cameron, 2003; Deignan, 2005; Semino, 2008). Perkara ini memberi gambaran jelas terhadap peranan Prosedur Pengenalpastian Metafora (*Metaphor Identification Procedure (MIP)*) iaitu kepentingan leksikal, pola dan variasi metafora menjadi nadi dalam kupasan metafora linguistik.

Dalam *CMT*, tuntutananya lebih spesifik bahawa metafora membolehkan manusia berfikir dan membincangkan soal pengalaman abstrak, kompleks, subjektif, kurang digambarkan, dan sering dikaitkan dengan sifat manusia. Hal ini menjadikan metafora sebagai fenomena linguistik dan kognitif penting serta menerangkan jumlah metafora perhatian yang telah diterima sejak tahun-tahun kebelakangan ini (Semino, 2008). Malah, ia juga telah mencadangkan metafora sebagai bahagian penting dalam kapasiti kreativiti dan inovasi yang membawa kepada perkembangan manusia moden dalam proses evolusi (Wong, 2017). Walau bagaimanapun, untuk menjelaskan penggunaan metafora dalam wacana, keperluan dalam mempertimbangkan pelbagai fungsi yang lebih spesifik yang boleh diperolehi dalam metafora (Pragglejaz, 2007). Perkara ini terpakai terutamanya apabila ungkapan metafora dipilih dalam banyak kemungkinan alternatif, atau apabila melibatkan kreativiti dan pola teks wacana (Halliday & Matthiessen, 2014; Pragglejaz, 2007; Semino, 2008).

Sehubungan dengan hubungan antara leksikal dan metafora, salah satu medium untuk memperoleh leksikal secara saintifik melalui kajian korpus linguistik, yang melihat bagaimana bahasa sebenarnya digunakan dalam konteks tertentu dan bagaimana boleh berbeza dari konteks ke konteks (Crawford & Csomay, 2016). Hubungan antara leksikal ini boleh dijelaskan dengan representatif senarai kekerapan





kata, kata kunci dan analisis penghubungan matriks yang memberikan dapatan konteks data yang autentik dan kajian empirikal. Sementara pemahaman variasi dan perbezaan kontekstual merupakan matlamat yang dikongsi oleh penyelidik dalam bidang penyelidikan linguistik yang lain, korpus linguistik menerangkan variasi bahasa dan penggunaan dengan melihat sejumlah besar teks yang dihasilkan dalam keadaan yang sama (Biber, Conrad, & Reppen, 2006). Hal ini dipanggil sebagai data raya (*big data*) di mana bentuk data maya dimunculkan hasil janaan alat statistik (perisian korpus linguistik). Konsep ‘*circumstance*’ atau ‘*context*’ atau ‘*situation*’ bergantung kepada bagaimana setiap peneliti menerangkan dapatan.

Korpus linguistik merupakan representatif kumpulan bahasa yang digunakan dalam komunikasi bagi penggunaan bahasa (Crawford & Csomay, 2016). Korpus linguistik juga menitikberatkan bersama kefahaman bagaimana manusia menggunakan bahasa dalam multikonteks (Suic, 2002). Analisis kumpulan bahasa ini memberikan satu representatif pola penggunaan bahasa secara semula jadi dengan memberi penekanan kepada leksikal yang disenaraikan, sama ada leksikal itu berbeza jenisnya atau menjadi kata kunci terhadap wacana yang digunakan. Pemerhatian daripada leksikal, pola dan konteks merupakan satu analisis korpus linguistik yang penting dan menjadi analisis ‘berpanduan korpus’. Sarjana korpus linguistik telah menggariskan definisi penyelidikan korpus linguistik, iaitu pertama, kajian empirikal, analisis pola sebenar yang digunakan dalam teks bahasa semula jadi; Keduaanya, ia menggunakan koleksi teks semula jadi yang besar dan berprinsip, yang dikenali sebagai analisis asas; Ketiganya, ia menggunakan komputer secara meluas untuk analisis, menggunakan





teknik automatik dan interaktif; Keempatnya, ia bergantung kepada teknik analisis kuantitatif dan kualitatif (Biber et al., 2006).

Senario statistik dalam korpus linguistik telah representatif data berbentuk nombor dan normalisasi terhadap leksikal yang muncul. Satu istilah yang boleh diketengahkan dan diguna pakai bagi merujuk statistik leksikal dan metafora adalah ketumpatan leksikal atau *lexical density*. Pendefinisian ketumpatan leksikal digunakan sebagai galas konsep statistik dalam korpus linguistik dan pengkomputeran (Gregori-Signes & Clavel-Arroitia, 2015). Pelbagai variasi leksikal yang muncul berdasarkan kekerapan yang disenaraikan menunjukkan nilai signifikan secara berstrata sama ada tinggi, sederhana atau rendah (*hapaxes*) (Muñoz-Espí & Bastida-Serra, 2019). Hal ini diselarikan dengan variasi metafora yang memiliki fitur ketumpatan metafora atau *metaphor density* dalam dapatan kebarangkalian metafora (*probabilites metaphor*) (Sardinha, 2011). Rasionalnya, penggunaan ketumpatan ini adalah rangkuman representatif statistik dalam menganalisis hubungan leksikal dan metafora sebagai pengukuran kekayaan leksikal (*measure of lexical richness*) terhadap data kajian (Johansson, 2008; Lu & Wang, 2017).

Fenomena penyelidikan metafora dan korpus linguistik bukan satu perkara baharu dalam dunia kajian global. Sebaliknya, matlamat serta metodologi menjadi nadi dalam kepelbagaian perspektif. Oleh itu, satu kajian yang meneliti manuskrip Melayu atau lebih dikenali dengan teks klasik diperhalusi, cerapan dengan menggunakan pandangan metafora menjadi landasan terhadap prinsip leksikal dalam teks bahasa tabii. Hal ini menjadikan, korpus linguistik sebagai landasan utama memberikan satu





impak yang empirikal dalam penyelidikan statistik deskriptif bahasa Melayu. Rasionalnya, keupayaan menterjemah segala ilmu perlu dipergiatkan dengan mengautomasikan kaedah dan analisis seterusnya memperoleh pola intelektual masyarakat Melayu dahulu dalam linguistik Melayu.

Sekilas, dalam konteks kajian seni bina atau alam bina keupayaan interpretasi trend atau fenomena industri terdahulu hanya digambarkan dengan kehadiran bangunan lama atau rumah tradisional yang merupakan sejarah visual objek yang masih kekal sebagai artifak. Sedangkan perkembangan teori seni bina kebanyakan bermula dengan nafas sarjana linguistik seperti Teori Semiotik Seni Bina berdasarkan Model Segitiga oleh Ogden-Richards, Semiotik atau Komunikasi Seni Bina oleh Ferdinand de Saussure, Teori Fenomenologi oleh Edmund Husserl (1859-1938), dan Teori Dekonstruktivisme oleh Jacques Derrida (Nik Lukman, 2018). Ahli falsafah dan linguistik memberi sumbangan yang besar dalam perkembangan seni bina pascamodenisme. Keupayaan menterjemah dan mengasah idea dalam konteks seni bina melahirkan reka bentuk bangunan yang dipacu berdasarkan keperluan masyarakat dunia. Sehubungan dengan itu, selari dengan konteks kajian ini menyentuh ilmu seni bina Melayu yang dirakam oleh catatan manuskrip Petua Membina Rumah. Kewajaran dalam meneliti manuskrip ini ialah perkembangan industri seni bina atau pembinaan dahulu perlu ditonjolkan sebagai sumber rujukan digital dan perolehan bahasa dalam laras seni bina. Di samping itu, dasar kerajaan dalam konteks seni bina kebangsaan dapat disokong dalam bentuk sumbangan warisan lama bertulis.





Berdasarkan sumber catatan bertulis ilmu seni bina, sarjana bukan sahaja mengolah dapatan falsafah dan kreatif pembangunan seni bina, sebaliknya peranan sains bangunan serta teknologi bangunan yang terdapat dalam manuskrip mampu dikaji. Hal ini demikian kerana metafora sebagai bahasa saintifik merupakan satu bentuk bahasa yang dibentuk berdasarkan pengalaman dan penghayatan sekeliling. Menurut Fadzllah Zaini & Nor Hashimah Jamaluddin (2019) wujud hubungan peribahasa Melayu serta manuskrip (*Malay Concordance Project*) dengan struktur seni bina rumah Melayu tradisional terhadap kehadiran unsur struktur seni bina. Hubungan seni bina dan metafora dapat memberi penjelasan kepada kognitif pemikir masyarakat Melayu dalam menonjolkan akal budi Melayu melalui ilmu seni bina. Pelbagai perkara boleh disentuh dalam konteks seni bina, manuskrip dan metafora. Leksikal yang terlibat sebagai metafora boleh dikaji secara saintifik. Seperti dalam subdisiplin sains bangunan menyentuh soal kerja tanah (*earthwork*), kerja pembersihan tapak (*site clearance*) faktor pengalir dan penambat bunyi dan lain-lain elemen yang dinyatakan dalam manuskrip Melayu. Di samping itu, terdapat juga elemen teknologi binaan yang boleh diamati seperti alatan pembinaan, struktur bangunan, tanggam (*jointing*) dan pelbagai lagi. Leksikal sedemikian dapat dicungkil dalam manuskrip Melayu. Hal ini disebabkan faktor manuskrip sebagai data akal Melayu yang boleh diteliti dengan baik dan bersumbangan.

Justeru, kajian ini bukan sahaja membahaskan ilmu manuskrip sebaliknya menonjolkan dapatan ilmu seni bina yang terdapat dalam manuskrip Petua Membina Rumah sebagai melihat kemajuan industri seni bina. Penggunaan analisis berpanduan korpus membawa satu perspektif terkini. Yang mana, memperincikan ketumpatan





leksikal yang dibawa dalam wacana teks klasik ini. Kemunculan leksikal tersebut boleh diteliti dengan prosedur korpus linguistik dan prosedur pengenalpastian metafora sebagai langkah petunjuk bentuk leksikal yang terdapat dalam manuskrip yang dipilih. Segala dapatan yang direpresentatitkan adalah berpandukan korpus sebagai pandangan makna yang berpaksikan prinsip korpus linguistik dan metafora ekspresi.

1.3 Permasalahan Kajian

Bahagian permasalahan kajian dibincangkan kepada tiga teras, iaitu teks klasik (manuskrip), metafora dan analisis berpandukan komputer. Tiga teras ini merupakan prinsip utama dalam penyelidikan ini, segala kelompongan, kewajaran dan perbandingan dilakukan untuk memperhalusi permasalahan kajian. Teras ini diselarikan dengan literatur (rujuk seksyen 2.4; 2.5; 2.6).

1.3.1 Kajian Teks Klasik (Manuskrip)

Penyelidikan ini telah memfokuskan manuskrip dan teks klasik sebagai data kajian (rujuk seksyen 3.6). Manuskrip atau teks klasik merupakan satu bentuk data bahasa tabii yang memuatkan segala bentuk leksikal, pemikiran, tatacara kehidupan, masyarakat dan hal ehwal kehidupan yang lain (Babeu, n.d.). Keupayaan kandungan





manuskrip dan teks klasik ini telah melangkaui satu tranformasi terhadap bentuk data digital. Terdapat beberapa penyelidikan yang telah dijalankan bagi mengupas teks klasik dalam pelbagai bahasa sumbernya. Kajian Al-Thubaity, (2015); Belinkov et al., (2019); Djellab et al., (2017); Jarrar et al., (2017); Oudah & Shaalan, (2017); El-Haj et al., (2015) dan Hammo et al., (2016) telah membangunkan korpus teks klasik atau teks lama bahasa Arab bagi penyelidikan laras agama dalam sosiolinguistik, impak *language-independent* dan *language-specific*, perbandingan bahasa standard dan dialek dan kajian sejarah bahasa dengan menggunakan skala data yang besar. Dalam pada itu, kajian-kajian teks klasik dalam bahasa lain giat diusahakan dalam penyelidikan. Kajian Koo, (2015); Nguyen, Le, Ho, & Nguyen, (2015); Aranberri, Labaka, Díaz de Ilarraza, & Sarasola, (2017); Rello, Baeza-Yates, & Llisterri, (2017); Timmis, (2018); Parada-Cabaleiro, Costantini, Batliner, Schmitt, & Schuller, (2019); Magistry, Ligozat, & Rosset, (2019); Abainia, (2019); dan Rubinstein, (2019). Perspektif kata pinjaman, struktur rajah pohon, penilaian terjemahan, kesalahan penulisan, analisis fonetik dan visual, pragmatik, kesan emosi, kelas kata, alih kod, dan sejarah bahasa. Penyelidikan sarjana global ini menunjukkan trend kajian teks klasik yang berhubungan dengan bahasa atau linguistik. Hal ini kerana, peranan teks klasik bukan sahaja dibincangkan atau dianalisis dalam kelompok kesusasteraan namun boleh dikaji dalam fungsi linguistik dan penggunaan dalam kandungan pendidikan bahasa. Fokus diberikan terhadap kajian korpus linguistik yang menggunakan data teks klasik.

Kontradiksi korpus linguistik dalam pemakaian data teks klasik adalah kajian berpanduan dan berasaskan korpus menghasilkan satu fenomena yang berbeza. Di mana data dipandu dalam mencari leksikal kandungan dengan mengaplikasikan analisis



berpandukan korpus. Berbeza dengan berasaskan korpus yang mencerpap data dengan aplikasi teori linguistik atau sebaliknya bergantung kepada fokus kajian. Hal yang sama turut berlaku, kajian Bingenheimer, (2015); A. Cook, (2018); Erjavec, (2015); Lu & Wang, (2017); dan List et al., (2017). Kajian teks klasik bahasa Cina banyak dijalankan menggunakan teori-teori linguistik dan masih menggunakan tatacara manual dalam analisis data. Kajian teks klasik kebanyakan melihat pembangunan, perubahan, penemuan, pencarian bahasa yang digunakan oleh masyarakat dahulu kemudiannya dibandingkan. Perbandingan ini bukan menjadi kayu pengukur terhadap kekurangan ilmu lama dan baharu, sebaliknya penambahbaikan perlu ada bagi mewujudkan kesempurnaan ilmu dalam bidang.

Di samping itu, teks ini merupakan karya yang kaya dengan sumber maklumat dan mempunyai ruang lingkup yang luas. Kajian Guy & Schwitter, (2017); Santamaría-García, (2017); Flöck & Geluykens, (2018); Mohamed & Oussalah, (2019); dan Norri, Junkkari, & Poranen, (2019) memerihalkan soal ketinggian nilai teks klasik dalam pelbagai ilmu. Kajian-kajian ini telah menyumbang dapatan kepada pembangunan ilmu secara berorganisasi. Keanekaragaman isi dan maklumat dalam teks klasik mempunyai keistimewaannya sendiri yang tidak terdapat dalam teks yang lain (Harun, 2006). Tambahan lagi, telah banyak juga penyelidikan umum teks klasik dijalankan bagi melihat fungsi teks. Kajian Birch, (2019); Cullen, (2019); Deps & Charlier, (2019); Kerzendorf, (2019); Li, Yang, & Lv, (2019); Moul, (2019); Namugera, Wesonga, & Jehopio, (2019); Raybould, (2019); Ricciardi, (2019); Simon et al., (2019); Spallone & Vitali, (2019); Stowe, (2019); Takahashi, (2019); Valdés Miyares, (2019); van Kranenburg, de Bruin, & Volk, (2019); Wang et al., (2019); dan Xu, Ritzhaupt, Tian,

& Umapathy,(2019) telah memaparkan sumbangan yang sarat dalam teks klasik. Antara kajian yang dilakukan adalah, fungsi teks dari segi visual klasik, budaya politik, era kesusasteraan, falsafah silang budaya, klasik dalam keagamaan, bio-perubatan, reka bentuk seni bina, astronomi, sains, muzik, naratif klasik, moral, sektor kerjaya, teks klinikal dan *Crouzon sindrom*. Oleh hal yang demikian, cerapan terhadap taburan leksikal perlu dilakukan bagi mendapat dapatan pelbagai bidang. Kebanyakan pengkaji masih menggunakan telaah manual yang memerlukan masa yang lama serta kebarangkalian leksikal disenyapkan akan tinggi kerana daya upaya secara mental mampu memberi bias kepada data teks. Salah satu medium yang boleh memproses data teks mentah ini melalui korpus linguistik yang menilai leksikal secara menyeluruh berdasarkan prinsipnya. Tambahan lagi, penjelasan secara statistik deskriptif yang menjelaskan kemunculan signifikan dengan memberikan bukti yang empirikal (Brezina, 2018). Kemunculan variasi leksikal, sama ada bersifat metafora atau literal sebagai fenomena kekayaan leksikal (Gregori-Signes & Clavel-Arroitia, 2015).

1.3.2 Kajian Metafora

Perspektif penyelidikan global, kajian metafora sentiasa bernafas baharu yang mempunyai kesinambungan antara penyelidikan. Kajian bahasa dan metafora ini merujuk kepada pengkaji bahasa seperti bahasa Inggeris, Cina, dan lain-lain yang telah mengkaji metafora. Kajian Yau, (2015); van Rij, (2016); Pavlinich, (2016); Curry & Chambers, (2017); Jing-Schmidt & Peng, (2017); dan Stanley, (2017) merupakan

kajian metafora dengan meneliti pendidikan bahasa, teks sastra lama dan wacana terkini. Kajian ini melihat representasi metafora yang muncul dalam teks dan mempengaruhi bidang yang diterokai para sarjana. Kelompongan dapat dilihat dengan implimentasi metafora linguistik pada tahap yang sederhana, sebaliknya peranan metafora konseptual telah melahirkan pelbagai bentuk andaian yang boleh disangkal berdasarkan kognitif pemikiran manusia yang sentiasa berubah. Pertimbangan terhadap penyelidikan metafora linguistik merupakan salah satu metodologi yang boleh diaplikasi dalam korpus linguistik. Hal ini kerana peranan korpus linguistik yang mampu menjelaskan representatif leksikal yang mempunyai rantaian terhadap metafora (Brugman, Burgers, & Vis, 2019). Kewajaran ini memberi nilai tambah terhadap kajian metafora dalam pelbagai bentuk wacana terutama dalam manuskrip serta teks klasik.

Keduanya, penyelidikan metafora dan korpus linguistik yang dibahagikan kepada tiga aspek utama, iaitu leksikogramartikal, metaforikal, budaya dan tekstual. Antara kajian yang terlibat dalam tiga aspek ini adalah kajian Devrim, (2015); Velázquez-Mendoza, (2015); Fontaine, (2017b, 2017a); García Marrugo, (2017); Gruzitis & Dannélls, (2017); Lu & Wang, (2017); Mey, (2017); V. J. van Rij, (2017); Wu, Lau, & Li, (2017) Reijnierse, Burgers, Krennmayr, & Steen,(2017); Scott, (2017); Purcell, (2018);Parada-Cabaleiro et al., (2019); Pizarro Pedraza, (2019); Tulloch, (2019); Ullmann, (2019); dan Guanio-Uluru, (2019). Kajian budaya dan tekstual ini merupakan satu perspektif analisis yang menggunakan metafora serta korpus sebagai medium utama dalam penelitian. Kajian ini berkenaan dengan keterangan atau kejelasan bahasa Inggeris dalam teks, refleksi meta-pragmatik, etnografi sejarah dan pertembungan budaya. Hal ini menunjukkan bahawa representasi metafora yang diteliti



dengan menggunakan korpus mampu menjelaskan budaya dan senario teks tersebut. Kajian metaforikal yang sentiasa mendedahkan interpretasi dalam korpus yang mampu menzahirkan pelbagai anggapan yang difahami dan dijelaskan kepada masyarakat.

Selain itu, kajian leksikogramartikal yang merupakan satu penyelidikan struktur metafora berdasarkan kelas kata yang digunakan. Struktur metafora ini merupakan pola kata yang digunakan dalam membentuk metafora. Leksikal ini membolehkan sarjana mengenal pasti metafora secara statistik di mana data itu boleh diuji kebolehpalsuan, keringkasan, kelengkapan dan keutuhan dalam membentuk metafora. Di samping itu, penggunaan Prosedur Pengenalpastian Metafora (*Metaphor Identification Procedure (MIP)*) sebagai kaedah eksplisit, boleh dipercayai, dan fleksibel untuk mengenal pasti kata-kata metafora yang digunakan dalam bahasa. Antara kajian yang telah dijalankan adalah literasi teks bahasa, neologism BREXIT, konteks tempatan, mati dan bunuh, serta kawalan NLP. Keupayaan dalam penelitian leksikogramartikal memperincikan lagi MIP dalam memberikan dapatan dan kewajaran dalam kajian metafora dan korpus linguistik.

Ketiganya, kajian metafora, korpus dan multidisiplin. Kajian Aubusson, Harrison, & Ritchie, (2006); Kloppmann-lambert, (2018); dan Ricciardi, (2019). Penyelidikan ini membawa metafora dan korpus kepada dapatan berbentuk multidisiplin (ilmu sains, teknologi dan seni bina). Kajian sedemikian amat kurang diambil perhatian kerana perspektif sarjana perlu bijak dalam mengunjurkan dapatan yang mengambil kira kepentingan pembangunan tamadun. Oleh itu, metafora yang disisipkan dalam teks dengan menggunakan korpus dan mampu dianalisis berpanduan





korpus bagi penelitian setiap leksikal yang muncul dan merepresentasikan metafora-lingusitik yang dikodkan dalam pelbagai teks bahasa.

1.3.3 Kajian Analisis Berpanduan Korpus

Pertama, pola penyelidikan analisis berpanduan korpus telah banyak dijalankan menerusi pelbagai korpus bahasa. Antaranya kajian berpanduan korpus penelitian terhadap bahasa Inggeris, Cina, Rusia, Sepanyol, Korea, Jerman dan Arab. Kesemua bahasa ini merupakan bahasa dunia yang sedang giat usaha dalam mendigitalkan data sebagai kemudahan penyelidikan sarjana dan pendidikan. Kajian Barteld, Biemann, & Zinsmeister, (2019); Belinkov et al., (2019); Cheung & Larson, (2018); W. R. A. Cook, (2017); Crosson, McKeown, Moore, & Ye, (2019); Gribanova, (2017); Ingham & Tubau, (2019); Jiménez-Zafra, Taulé, Martín-Valdivia, Ureña-López, & Martí, (2018); Kimmelman, Pfau, & Aboh, (2019); O'Brien, Mohamed, Yussof, & Ng, (2019); Rahimi & Zhang, (2019); Roehrs & Sapp, (2018); Sepúlveda-Torres, Sanches Duran, & Aluísio, (2019); Sun & Zhang, (2018); Yao, (2018); Yeung, Ho, Chan, & Chung, (2019); dan Yoon, (2018) merupakan satu fenomena penyelidikan analisis berpanduan korpus dalam pelbagai data teks bahasa sebenar. Namun begitu, di Malaysia kesedaran dalam era industri 4.0 data digital sedemikian menjadi keperluan utama. Pengkaji perlu menjadi bijak dalam mengelola data berbentuk digital. Hal ini dapat memartabatkan bahasa Melayu bertujuan mengumpul segala bentuk penulisan dan lisan sebagai hub simpanan data bahasa, di samping menjaga nilai ilmu manuskrip Melayu. Kepentingan





data raya (*bigdata*) sebagai kumpulan variasi leksikal yang boleh membantu guru untuk pengajaran dan pembelajaran bahasa sebagai satu langkah memperkayakan leksikal mental pelajar (McEnery & Xiao, 2011).

Seterusnya, kajian penelitian korpus linguistik yang menggunakan teori linguistik sebagai asas penyelidikan korpus. Kajian sedemikian telah menetapkan bentuk dapatan berdasarkan teori yang digunakan. Rujukan dan penambahbaikan daripada kajian lalu menjadi rantaian perbahasan dalam kajian-kajian sedemikian. Antara kertas penyelidikan adalah kajian Alfuraih, (2019); Allen, Likens, & McNamara, (2019); Bar-On & Kuperman, (2019); Barteld et al., (2019); Beauvais, (2019); Beck, (2019); Bruening, (2019); Dal Maso & Giraudo, (2019); Ingham & Tubau, (2019); dan Rago, Marcos, & Diaz-Pace, (2018). Namun, bentuk kajian analisis berpandukan korpus memerlukan bentuk teori atau prosedur yang fleksibel bagi menjelaskan senario teks. Keperluan ini penting sebagai landasan untuk memerihalkan pola intelektual masyarakat terdahulu serta variasi leksikal yang menjadi perhatian utama. Kekurangan daripada sudut analisis dengan melihat fenomena leksikal daripada kaedah statistik deskriptif iaitu ketumpatan leksikal jarang dibahaskan. Sebaliknya pengkaji terdahulu mengambil perspektif konkordans dalam teks yang dianalisis. Ketumpatan leksikal ini penting bagi penilaian prinsip realistik, keserasian (*kongruen*) dan *replicability* dalam analisis berpandukan korpus. Ketumpatan leksikal dalam kajian ini meliputi taburan kata dan perkataan yang digunakan metafora (MUW) dianalisis dengan pengujian skor signifikan yang ditetapkan iaitu mi-skor, t-skor dan logDice.



Lain daripada itu, kajian berkenaan dengan berpandukan korpus dengan menggunakan data teks lama atau teks klasik. Antara kajiannya adalah Bratishenko, (2019); Drackley, (2019); Kokkola, (2019); Liang & Zheng, (2018); Neidorf, Zhao, & Yu, (2019); Rainer, (2018); Roehrs & Sapp, (2018); Rutten, Krogull, & Schoemaker, (2019); dan Spotti, Kroon, & Li, (2019). Kajian ini mewajarkan kajian teks lama boleh dilakukan dengan analisis korpus linguistik. Gambaran penyelidikan ini memfokuskan kepada teks kesusasteraan lama yang diambil dan dikaji semula bagi melihat tiga perkara penting iaitu bahasa, pemikiran dan tamadun. Apabila menyentuh soal teks klasik atau manuskrip lama, kelompangan yang jelas adalah kajian teks klasik atau manuskrip Melayu menggunakan analisis korpus linguistik. Pandangan holistik sedemikian menjadi keperluan utama bagi penyelidikan terhadap teks klasik Petua Membina Rumah. Hal ini dapat memberi gambaran yang jelas terhadap leksikal dan metafora seni bina yang di representatif dalam wacana tersebut.

1.3.4 Fenomena di Malaysia Kajian Teks Klasik-Metafora-Analisis Berpandukan Korpus

Berdasarkan kajian teks klasik, metafora dan analisis berpandukan korpus, secara terperinci keperluan untuk melaksanakan kajian ini adalah signifikan. Keperluan daripada sudut data teks klasik, iaitu manuskrip Melayu tidak banyak dikaji dalam linguistik dengan mengaplikasikan metafora ekspresi. Penggunaan analisis berpandukan korpus dalam penilaian metafora ekspresi juga pada tahap sederhana.



Sehubungan dengan kajian lepas ini, kajian ini mengkaji hubungan ketumpatan antara leksikal dan metafora berdasarkan analisis berpandukan korpus mengguna pakai prosedur pengenalpastian metafora dapat memberikan satu perspektif baharu terhadap kajian teks klasik Melayu.

Pertama, fokus diberikan terhadap data kajian iaitu teks klasik atau manuskrip Melayu. Hal ini kerana, tautan perbincangan berteraskan ilmu korpus linguistik dan metafora. Oleh hal yang demikian, perspektif teks klasik atau manuskrip perlu dikupas berdasarkan kandungan dan bukan pada peringkat bandingan kesusasteraan. Kajian manuskrip perlu dijalankan daripada perspektif kandungan disiplin supaya memperoleh istilah serta pemakaian bahasa Melayu terdahulu (Shaharir, 2000). Penyelidikan Ahmad, Lukman, & Yusof, (2016); Amer Hudhaifah, (2017); Faisal & Wahidah, (2012); Filzah, Mat, Wan, & Firdaus, (2018); Jelani, (2015); Muhammad Akmal R.Azmi & Faudzinain, (2018); Riswadi & Mustaffa, (2017); Wan Mohd Dasuki & Radziah, (2014); Ab. Razak, (2005); Ab. Razak, Pramono, & Haslan, (2016) ; Abdullah, (1992, 1993); Amat Juhari, (2000); Fairuzah, Salmah, & Mohd Syukri, (2013); Gallop, (2013); Hashim, (1993); Khairoslinda, Zubir, & Yusmilayati, (2017); Mohamad Zain, (2005); Muhamad Shafiq, (2016); Norhasnira & Faisal, (2020b, 2020a); Raja Masittah, (1993); dan Adi Yasran et al., (2016). Kebanyakan kajian sedemikian membincangkan kajian teks klasik atau manuskrip bukan sastera. Bukan sastera merupakan satu bentuk penulisan manuskrip seperti petua, kitab perubatan, gempu, warkah dan pelbagai lagi. Berdasarkan kenyataan Ding Choo Ming (2016) masih banyak ilmu dan maklumat peribumi Melayu dari pelbagai bidang belum diketahui.



Maklumat ini perlu disampaikan dan dimanfaatkan bagi menjelaskan senario keilmuan terdahulu yang dipadam berdasarkan kurang rantaian serta kesinambungan dalam mewarisi perkara ini. Kajian Nurhayati, (2008); Rofa & Mohd, (2010); Shaharir & Affandi, (2019); Ming, Chambert-Loir, & Pudjiastuti, (2014); Nazri & Abdul Latif, (2018); Nor Hashimah, (2013); dan Yakob, (2018) turut menjelaskan dapatan yang berguna seterusnya saling memberi manfaat dalam masyarakat terutama konteks terkini dan realistik. Variasi leksikal terdahulu sangat memberi manfaat dalam pengeklakan bahasa atau pembentukan istilah (Juhaida, Khairuddin, Faidzul, & Zamri, 2013). Manuskrip ialah data bank bertulis tentang akal budi, citra, intelek, rasa, dan kebijaksanaan orang Melayu sejak zaman silam. Oleh itu, data bank bertulis perlu diberikan nafas terkini dengan pembangunan pangkalan data korpus yang lebih terperinci dan mudah dijana. Kerancangan pangkalan data korpus ini bermula dengan pembentukan Korpus DBP, kemudian korpus khusus *Malay Concordances Project* yang diusahakan pihak Universiti Kebangsaan Australia (ANU). Namun, bagi pemilikan data raya (*bigdata*) manuskrip masih kurang di Malaysia pendedahan bagi pembangunan korpus ini masih sederhana. Kajian ini merupakan satu usaha ke arah mendigitalkan data tersebut.

Hasil buku Al-Ahmadi (2016) telah menerbitkan Petua Membina Rumah dari perspektif antropologi dan dibandingkan dengan industri pembinaan Nusantara. Perspektif kajian tersebut memfokuskan budaya dan sosial masyarakat dalam pembinaan rumah. Namun, tidak menjelaskan hubungan leksikal atau metafora ekspresi yang wujud. Lanjutan itu, kajian ini memperincikan dari sudut korpus linguistik yang membincangkan soal leksikal dan kemunculan metafora linguistik



dalam Petua Membina Rumah. Kajian Khairudin, Salleh, & Aripin (2018) telah melihat faktor pembinaan rumah diselarikan dengan masa (bulan Islam) dirujuk berdasarkan manuskrip Petua Membina Rumah. Rasional industri terdahulu dalam menyemak keadaan masa, bulan dan pengamatan sekeliling dalam menjalankan kerja-kerja mendirikan rumah. Oleh hal yang demikian, eksplorasi dalam taburan leksikal perlu diperincikan bagi menjelaskan hubungan leksikal dan metafora sebagai sesuatu ilmu rumah terdahulu yang masih relevan dan realistik. Kekurangan dalam mencerap ilmu seni bina Melayu dalam teks ini, ramai pengkaji menyifatkan bahawa tiada bahan atau data yang perlu diteliti dan menerangkan soal industri seni bina terdahulu.

Seterusnya, kajian yang melibatkan metafora dan manuskrip seperti ilmu seni bina ialah kajian Fadzllah Zaini & Nor Hashimah Jamaluddin (2019). Kajian ini meneliti unsur seni bina dalam peribahasa Melayu dan digarapkan bersama data teks klasik dalam *Malay Concordance Project*. Kajian metafora konseptual yang memberi rangsangan kepada interpretasi maklumat peribahasa terutama istilah seni bina dan menjelaskan akal budi Melayu sebagai dasar Semantik Inkuisitif. Sama seperti kajian Mohd Rashid, (2012), Suriati Zakaria, (2016), Nirwana Sudirman, Zulkifley Hamid, Masitah Mad Daud, & Nik Zulaiha Zakaria, (2017), Asad & Nor Hashimah, (2018); Julaina, Nor Hashimah, & Junaini, (2017, 2018); dan Nur Afiah & Nor Hashimah, (2016). Namun dalam kajian ini, perubahan landasan diberikan kepada prosedur pengenaltastian metafora, iaitu sebagai metafora linguistik. Seperti kajian Ramanathan, Hoon, & Paramasivam, (2018) yang telah mengaplikasikan prosedur pengenaltastian metafora dalam data media sosial *twitter* politik semasa pilihan raya di Malaysia. Namun, penggunaan prosedur ini pada peringkat pengenaltastian metafora





secara penandaan kewujudan metafora tanpa merujuk kemunculan nilai pengujian skor dalam teks dengan perisian korpus linguistik (rujuk seksyen 3.4).

Namun hal ini tidak berlaku dalam senario kajian korpus linguistik di Malaysia. Kajian seperti Imran Ho-Abdullah (2011); Simah Mat Awal, Azhar Jaludin, Anis Che Abdul Rahman, & Imran Ho-Abdullah (2019) Normi, (2020); dan Sabrina, Fariza, Azhar, & Nadiah, (2020). Kajian ini telah mengambil ruang penyelidikan secara kuantitatif dan kualitatif. Kebanyakan kajian menjurus kepada kualitatif sebagai menerangkan konteks ayat bagi sesebuah kata kunci iaitu konkordans. Bagi kajian representatif leksikal secara statistik masih kurang diterokai. Keperluan instrumen dalam menyentuh statistik leksikal kurang didedahkan. Justeru, instrumen kajian ini dipilih berdasarkan kepakaran korpus linguistik di Universiti Lancaster telah membangunkan perisian LanksBox 4.5 sebagai satu bentuk baru alat statistik untuk mengkaji leksikal khususnya. Kajian seperti Gablasova, Brezina, & Mcenery, (2017); dan Mc Enery, Brezina, Gablasova, & Banerjee, (2019) telah menunjukkan jaringan leksikal bagi membantu pelajar bahasa untuk memperkayakan perbendaharaan kata dalam ruang mental.

Kesinambungan sebagai keutamaan perbendaharaan kata, perkara sedemikian mampu membantu dalam menyelesaikan inventori pada tahap pemakaian leksikal dan konteks realistik terdahulu. Ini kerana kajian Farhana, Faizah, & Shah, (2015) telah menjalankan penyelidikan mengenai inventori bangunan warisan yang memerlukan istilah dan makna yang digunakan bagi setiap struktur, bahan, reka bentuk dan aspek yang lain. Terdapat beberapa maklumat telah membatasi perolehan istilah yang sesuai



memberi gambaran terhadap fizikal seni bina terdahulu. Keupayaan metafora dan data korpus ini dapat membantu memperoleh pelbagai maklumat leksikal dalam kelas kata di samping membentuk makna sebagai keperluan penyelarasan istilah mengikut laras industri.

Tegasnya, kajian ini mengkaji hubungan antara leksikal dan metafora dalam manuskrip Melayu Seni Bina bagi menutup dan meminimumkan kelompangan kajian teks klasik atau manuskrip, metafora dan korpus linguistik. Secara langsung mewujudkan satu perspektif baharu terhadap kajian multidisiplin yang telah mempertimbangkan keperluan antara bidang khususnya dalam pendidikan bahasa Melayu. Pada amnya, implementasi data raya sebagai satu revolusi industri pendidikan

4.0.

1.4 Objektif Kajian

Objektif kajian ini merupakan hasil daripada permasalahan kajian yang telah dibincangkan. Terdapat dua fokus utama dalam objektif kajian ini, iaitu korpus linguistik dan metafora yang dianalisis berpandukan korpus khusus, iaitu Korpus Petua Membina Rumah.



Objektif pertama (O1) adalah kajian ini mengenal pasti ketumpatan leksikal (kekerapan dan taburan kata) dalam data korpus. Taburan kata memiliki dua analisis, iaitu kekerapan kata (*wordlist*) dan kata kunci (*keywords*). Janaan data dilakukan sepenuhnya menggunakan perisian *#LancsBox 4.5* merupakan kaedah statistik dalam korpus linguistik yang diasaskan oleh Brezina, (2018). Representasi ketumpatan leksikal dapat dijelaskan dan dikaitkan berpandukan konteks leksikal (*KWIC*).

Hasil daripada representasi ketumpatan leksikal tersebut, **objektif kedua (O2)** memfokuskan kepada menganalisis penghubungan matriks (*co-occurrence matrik*) antara leksikal. Peringkat ini menggunakan analisis kolokasi antara nod yang menghasilkan representatif graf visualisasi geomatriks leksikal berdasarkan *GraphColl*. Terdapat satu skor yang dipilih dalam penentuan nilai signifikan, iaitu logDice.

Setelah itu, peranan **objektif ketiga (O3)** sebagai kesinambungan analisis penghubungan matriks, iaitu menganalisis leksikogramartikal metafora yang terdapat dalam data korpus. Ketumpatan leksikal melalui senarai kata kandungan yang memiliki nilai signifikan dipilih bagi melihat pola metafora grammatikal. Pada fasa analisis ini tiga skor dipilih bagi menilai signifikan kelas leksikal iaitu mi-skor, t-skor dan logDice. Kesejajaran itu, analisis kebarangkalian Metafora juga digunakan bagi meneliti leksikal yang digunakan sebagai metafora. Kebarangkalian ini menggunakan skor *Metaphor Used Word (MUW)* diselarikan dengan kerangka Sardinha (2011). Pada fasa ini, kemunculan ketumpatan metafora (*metaphor density*) digunakan untuk representatif dapatan dimensi variasi metafora.



Sehubungan representatif ketumpatan metafora (*metaphor density*), **objektif keempat (O4)** adalah menghuraikan dimensi variasi metafora berdasarkan pemboleh ubah metafora iaitu topik dan sarana. Variasi ini menjelaskan peranan leksikal metafora dalam senario teks manuskrip Melayu yang telah didigitalkan dalam data korpus. Dalam huraian ini, fokus diberikan kepada KWIC dan konteks yang terlibat.

1.5 Soalan Kajian

Terdapat empat soalan kajian dalam penyelidikan Korpus Petua Membina Rumah (KPMR), iaitu;

1.5.1 Taburan dan kekerapan kata merupakan satu kelompok leksikal yang terdapat dalam Korpus Petua Membina Rumah. Apakah bentuk dan ketumpatan leksikal yang berlaku berdasarkan taburan dan kekerapan kata?

1.5.2 Nilai matriks penghubungan antara leksikal adalah satu ko-kejadian (*co-occurrence*) yang sering berlaku dalam Korpus Petua Membina Rumah. Bagaimanakah pola nilai kolokasi penghubungan matriks yang memaparkan signifikan leksikal?

1.5.3 Kepentingan leksikogramatikal dan kebarangkalian metafora sangat penting bagi menjelaskan perkataan yang digunakan sebagai metafora dalam Korpus

Petua Membina Rumah. Bagaimanakah pola leksikogramatikal dan kebarangkalian metafora yang terdapat dalam teks korpus?

1.5.4 Dimensi variasi metafora menjelaskan metafora linguistik dan ekspresi metafora dalam Korpus Petua Membina Rumah. Apakah variasi metafora yang dikodkan sebagai metafora linguistik Melayu dalam teks korpus?

1.6 Kepentingan Kajian

Kajian ini merupakan satu penyelidikan ketumpatan leksikal dalam manuskrip Melayu, fokus diberikan kepada tiga kepentingan utama, iaitu teks klasik (manuskrip), metafora dan korpus linguistik. Penyelidikan ini merupakan satu paradigma dalam penyelidikan teks klasik (manuskrip Melayu) yang disuntik analisis metafora dan korpus linguistik sebagai metodologi. Penyelidikan saban tahun bertukar-tukar medannya dan dapatan sentiasa berubah kerana kemajuan dalam kajian berlaku merentasi kelompangan dalam literatur. Maka, signifikan kajian ini boleh dibahagikan kepada beberapa perkara, iaitu kemajuan ilmu dalam teks klasik, kesejajaran analisis berpandukan korpus terhadap prosedur pengenalpastian metafora, nilai tambah literatur dalam linguistik Bahasa Melayu, transformasi keterbacaan manual kepada automatik serta representatif ilmu Seni Bina sebagai manfaat dalam bidang Sains dan Teknologi Bangunan. Perbincangan lanjut dikupas seperti berikut;



1.6.1 Kemajuan Ilmu Dalam Teks Klasik

Apabila memerihalkan teks klasik (manuskrip) Melayu, gambaran yang luas diberikan terhadap kemajuan masyarakat terdahulu dengan hadirnya kerajaan-kerajaan Melayu yang aman dan kadang kala bergolak, tetapi juga diperihalkan maklumat praktikal dan moral lain tentang kemandirian serta peningkatan taraf hidup orang Melayu. Berlatarbelakangkan persekitaran fizikal, sosial, ekonomi, politik, nilai, kepercayaan, kebijaksanaan dan fikiran mereka, teks-teks tersebut telah memberikan maklumat yang baik, malah representatif tentang masyarakat Melayu. Ribuan manuskrip Melayu yang masih penting kepada masyarakat Melayu masa kini, dan masa depan sebagai teras maklumat dan pendidikan (Ding Choo Ming, 2016). Malah, tidak kira maklumat tersebut bersifat sastera, separuh-sastera atau bukan sastera, disampaikan secara eksplisit atau implisit, semua manuskrip ini dipandang sebagai teras penting dalam masyarakat. Hal ini demikian kerana, perbincangan meliputi hal ehwal orang, budi, institusi sosial, intelek, budaya, ekonomi, politik, sains, teknologi, pertanian, penujuman, citra rasa, pendidikan, hiburan, dan perubatan.

Semua teras ini penting untuk menjamin penerusan jati diri orang Melayu masa kini dan masa depan (Hassan, 2016). Penyelidikan mesti diteruskan kerana hanya 30% daripada manuskrip yang berjumlah 10,000 judul dengan salinannya itu sudah dikaji, bagi mentafsir semula tanggapan sarjana Eropah yang mengkaji manuskrip Melayu lebih daripada perspektif sejarah (Ding Choo Ming, 2016). Dalam pada itu, banyak penemuan yang boleh memberi manfaat dengan penyelidikan melibatkan multidisiplin. Walaupun pada hakikatnya, tempat manuskrip adalah di muzium dan perpustakaan,





namun isinya harus dimanfaatkan dengan memindahkan akal budi, ilmu dan nilainya ke dalam jiwa masyarakat dan manuskrip tidak wajar dipandang sebagai artifak, bahan mati, penghias muzium dan perpustakaan semata-mata.

Oleh itu, kajian ini signifikan bagi melihat kepelbagaian khazanah ilmu terutama dalam variasi leksikal. Pemakaian leksikal oleh sarjana terdahulu telah memberikan satu bentuk ilmu berdasarkan hasil wacana ini. Pola susunan leksikal sehingga membentuk wacana manuskrip Melayu ini perlu dilihat sebagai representatif tradisi tulisan terdahulu. Penting juga dilihat struktur pemikiran Melayu melalui pola hubungan leksikal yang diguna pakai. Pengulangan atau pembentukan variasi leksikal baru merupakan bentuk wacana yang baik kerana kekayaan leksikal perlu diukur selari keupayaan ilmu masyarakat Melayu terdahulu. Manuskrip atau teks klasik bukan medium kuno yang perlu dibekukan sebaliknya sarjana perlu membedah setiap inci leksikal yang digunakan sebagai bentuk aktiviti kebahasaan. Perkara ini dapat mengekalkan variasi leksikal yang kian pupus terutama mengikut sektor pembinaan negara. Seperti ekonomi dan kepelbagaian industri serta pentadbiran kerajaan.

1.6.2 Kesejajaran Analisis Berpanduan Korpus Terhadap Prosedur Pengenalpastian Metafora

Perkara ini merupakan satu mauduk dalam kajian ini, perihal analisis berpanduan korpus yang diaplikasikan bersama Prosedur Pengenalpastian Metafora (MIP) oleh



Pragglejaz (2007). Menurut salah seorang ahli dalam kumpulan Pragglejaz telah memberi kefahaman terhadap pentingnya para sarjana dalam menggunakan prosedur ini. Berdasarkan Semino (2008), aplikasi prosedur ini memberikan tiga perkara utama iaitu metafora linguistik, metafora ekspresi dan perkataan yang digunakan metafora (MUW). Kepentingan MIP dilihat menerusi tiga perkara tersebut, yang memerihalkan rangsangan leksikal dalam menjadi metafora. Perubahan makna leksikal berlaku apabila hubungan antara leksikal memberikan kesan tinggalan makna sehingga mewujudkan metafora ditandai sebagai MUW. Oleh hal yang demikian, kesejajaran ini dilihat kepentingan representatif yang dinilai menerusi analisis berpandukan korpus yang melihat leksikal kerap dan kolokasi (bentuk hubungan) digunakan dalam MIP. Hal ini menunjukkan satu representatif terhadap variasi kefungsian leksikal yang terkandung dalam teks.

Apabila disejajarkan bersama analisis berpandukan korpus, prosedur ini memberikan impak positif dengan melibatkan nilai signifikan dalam bantuan menjelaskan konteks dan leksikal yang terkandung dalam sesebuah penyelidikan metafora bersama korpus linguistik. Hasil tindakan ini, telah membantu banyak aspek bagi menemukan titik kefahaman dalam pola wacana yang pelbagai. Kepentingan dalam linguistik bahasa Melayu merupakan keupayaan prosedur serta metodologi sedemikian dalam menyentuh soal leksikal, kelas kata, hubungan matriks leksikal dan representatif leksikal yang dibentuk seterusnya berlaku dalam teks klasik Melayu. Dalam usaha mencari keupayaan kebijaksanaan masyarakat terdahulu, para sarjana tidak perlu mencari orang lama sebaliknya mengkaji segala bentuk manuskrip yang disimpan rapi (Hernández-Campoy & Conde-Silvestre, 2012). Maka, telaah leksikal



manuskrip dengan pelbagai bentuk larasnya sama ada metafora atau tidak perlu dilakukan kerana setiap yang dicatatkan merupakan satu pengalaman pengarang yang mencerminkan pola intelektual masyarakat (Ding Choo Ming, 2016; Nor Hashimah, 2013).

Oleh hal yang demikian, kepentingan kajian ini yang sejajarkan analisis berpandukan korpus dan prosedur pengenalanpastian metafora (MIP) merupakan bentuk penyelidikan menarik hubungan antara leksikal dan metafora dalam teks klasik atau manuskrip Melayu. Hubungan ini dilihat dengan adanya pola ko-kejadian leksikal sehingga digunakan sebagai metafora. Kajian ini tidak mewujudkan satu metafora baharu sebaliknya mengekalkan bentuk metafora yang sudah sedia dalam teks klasik ini. Berdasarkan variasi leksikal yang dibentuk dalam pola, kemudian ditandakan sebagai metafora Melayu. Rasionalnya, impak pemikiran sekarang tidak mengganggu pemakaian makna dalam teks klasik Melayu. Ini akan menjamin penandaan metafora hasil daripada analisis berpandukan korpus dan metafora.

1.6.3 Nilai Tambah Literatur dalam Linguistik Bahasa Melayu

Saban tahun penyelidikan ilmu bertambah, para sarjana berusaha dalam memberikan dapatan yang terbaik dan sentiasa dibahaskan melalui literatur. Berdasarkan kajian ini yang menyentuh tiga keutamaan iaitu manuskrip, metafora dan korpus linguistik telah membuka lembaran penyelidikan statistik dalam bahasa. Penggunaan teks manuskrip





yang dibangunkan sebagai data korpus memberi kemudahan untuk mencerp lekikal berpandukan ketumpatan yang relevan. Kajian ini menyentuh aspek ketumpatan lekikal melalui analisis taburan kata dan analisis pengenapastian metafora. Ketumpatan yang diklasifikasikan sebagai nilai signifikan dalam ko-kejadian dalam senario teks manuskrip Melayu. Kajian ini juga memerikan senario lekikal yang telah dikodkan dalam linguistik Melayu terutama metafora.

Menurut Imran Ho-Abdullah, (2011), penerokaan data kekerapan melalui kaedah statistik dalam paradigma kajian linguistik dan bahasa berpandukan korpus dapat memberikan petunjuk serta pewajaran kepada persoalan lanjutan yang menuntut penyelesaian dalam pemantapan dan pembinaan ilmu linguistik. Tambahan juga Brezina (2018) menyatakan perubahan kaedah statistik perlu diperbaharui dari masa ke masa untuk mendapatkan satu dapatan yang baik, penambahan dalam representatif jaringan kolokasi dapat menentukan pola ko-kejadian yang berlaku dalam sesebuah teks. Hal ini sangat penting dan boleh diguna pakai dalam analisis linguistik bahasa Melayu bukan untuk menunjukkan kemajuan analisis akan tetapi kerelevanan kajian masa kini.

Kajian ini amat signifikan kerana mewujudkan satu perspektif penyelidikan korpus linguistik di Malaysia dengan melibatkan statistik sebagai penentuan hubungan lekikal dan metafora. Kekerapan lekikal yang terlibat dianalisis bagi mendapatkan ketumpatan lekikal dan variasi pola ko-kejadian lekikal yang terlibat. Kolaboratif antara lekikal menghasilkan bentuk metafora Melayu dalam wacana yang kompleks. Perbincangan statistik lekikal jarang sekali dibincangkan pada peringkat tesis





mahupun jurnal di Malaysia. Statistik leksikal ini merupakan bentuk penyelidikan yang diketengahkan bagi menangani masalah istilah yang sukar diperolehi. Penilaian leksikal sedemikian, mampu memberikan dapatan variasi leksikal. Di samping itu, pemakaian leksikal dalam KWIC sebagai contoh konteks atau ayat. Kesepaduan ini memberikan nafas dalam dunia penyelidikan bahasa untuk kajian statistik bahasa berteraskan teks terutama teks klasik Melayu.

1.6.4 Transformasi Keterbacaan Manual kepada Automatik

Perkembangan teknologi komputer telah dihidupkan kembali menerusi bidang Korpus linguistik. Keterbacaan mesin adalah sifat hakikat korpus moden. Hal ini disebabkan penggunaan korpus berkomputer, dan program komputer yang memudahkan analisis linguistik, yang membezakan korpus elektronik moden dari korpus awal '*drawer-cum-slip*' (Biber et al., 2006). Korpus berkomputer boleh diproses dan dimanipulasi dengan cepat pada kos yang minimum. Misalnya mencari, memilih, menyusun dan memformatkan. Komputer juga memproses data yang mudah dibaca mesin dengan tepat dan konsisten. Di samping itu, komputer dapat mengelakkan kecenderungan (*bais*) manusia dalam analisis, sehingga menjadikan hasilnya boleh dipercayai (Renouf, 1993). Kebolehbacaan mesin membolehkan pemprosesan automatik selanjutnya dilakukan pada korpus supaya teks korpus dapat diperkayakan dengan pelbagai metadata dan analisis linguistik (Crawford & Csomay, 2016).



Oleh itu, dalam transformasi ini analisis korpus linguistik mampu memerikan satu kepuasan dalam meneliti segala bentuk wacana yang bergantung kepada batasan sesebuah kajian. Analisis pada kadar automatik memberikan satu perspektif yang baik dan menggalakkan kajian yang melibatkan data mentah serta semula jadi. Kebolehbacaan data bahasa semula jadi ini mampu memperincikan leksikal yang terkandung dalam teks tersebut. Fenomena ini merupakan satu nilai positif dalam kajian bahasa dan linguistik. Penggunaan perisian komputer ini, terutamanya dalam kajian ini yang telah memilih *#LancsBox 4.5* memberikan satu instrumen baharu dalam kaedah statistik linguistik berkomputer. Perisian ini merupakan satu instrumen baharu dengan mengambil kira keseluruhan data teks untuk dianalisis.

1.6.5 Representatif Ilmu Seni Bina : Sebagai Manfaat dalam Sains dan Teknologi Bangunan

“Ilmu tidak pernah tua, sebaliknya menjadikan minda sentiasa muda” merupakan kata pujangga yang tersisip di laman Pusat Manuskrip Melayu Kebangsaan. Sememangnya apa yang terkandung dalam teks klasik ini sangat lama dan berusia. Namun, disebalik ilmu yang terkandung ialah khazanah budaya masyarakat yang diperhalusi sepanjang zaman. Apabila diteliti ilmu seni bina Melayu, rata-rata sarjana melihat artifak objek/bangunan secara langsung tetapi selakan manuskrip lama hanya ditelaah dalam perspektif budaya atau tidak disentuh. Bagaimana orang Melayu dahulu mendirikan rumah? Apakah mekanisme yang digunakan untuk projek yang dijalankan? Pelbagai persoalan mampu dikodkan jawapan dalam manuskrip Melayu seni bina. Ilmu kuno ini



bukan ketinggalan zaman sebaliknya berlaku dalam kemuncak kemodenan masa kini. Penelitian ilmu seni bina melalui leksikal ini penting dan menyumbang kepada semua bahagia daripada pendidikan, istilah, reka bentuk, identiti kebangsaan dan pelbagai lagi.

Kupasan aspek kandungan teks klasik Petua Membina Rumah mempunyai satu representatif yang tersendiri. Kehadiran pelbagai ilmu sains dan teknologi bangunan disisipkan dalam manuskrip tersebut. Cabang ilmu seni bina telah dibahagikan kepada beberapa subdisiplin, iaitu sains bangunan dan teknologi bangunan. Sains bangunan diklasifikasikan sebagai ilmu sains yang digunakan dalam pembinaan bangunan seperti kajian bunyi, tanah, angin, cahaya, bahan dan lain-lain. Manakala dalam konsep teknologi binaan bangunan lebih memfokuskan kepada teknologi yang digunakan dalam binaan bangunan seperti tanggam, struktur dan lain-lain. Perihal sedemikian berlaku dalam manuskrip petua membina rumah. Namun, perlu penelitian yang khusus dan terjemahan kepada kewajaran dalam pemilihan objek tersebut. Manuskrip sebagai sumber bertulis dan lisan yang melambangkan pengalaman rakyat mampu menterjemahkan elemen yang digunakan dalam seni bina Melayu. Sokongan daripada kajian Fróes & Fróes (2019) dan Onians (1992) menyatakan hubungan metafora dalam menterjemahkan seni bina mampu melahirkan kepentingan terhadap fenomena makna yang implisit.

Oleh yang demikian, dalam kajian ini memperoleh dua perkara utama, iaitu ilmu seni bina berdasarkan leksikal yang dipakai dan ilmu seni bina yang disisipkan secara tersirat melalui metafora. Kedua-dua perkara ini penting kerana kehadiran wacana yang kompleks ini telah menekankan konsep penulisan masyarakat Melayu yang dibentuk



berdasarkan persekitaran mereka. Kekompleksan ini telah memberi penekanan terhadap kesukaran dalam memahami teks klasik Melayu dan boleh diatasi melalui penelitian leksikal secara terperinci. Manfaat ini memberikan satu perspektif baru dalam metodologi mengkaji keupayaan leksikal yang digunakan dalam teks mengikut laras yang terlibat. Hasil ketumpatan leksikal dan metafora ini menjawab kepada senario teks klasik Melayu (manuskrip).

1.7 Definisi Operasional

Pendefinisian istilah dan konsep yang digunakan dalam kajian ini, khususnya melibatkan teori, konsep dan data yang dijelaskan sebagai panduan kajian.

1.7.1 Teks Klasik (Manuskrip)

Teks Melayu Klasik ditakrifkan sebagai bentuk teks Melayu bertulis yang telah digunakan di seluruh dunia Melayu di antara abad keempat belas dan abad kesembilan belas (Chambert-Loir, 2018). Menurut Shahrir (2000), teks klasik Melayu merupakan catatan ilmu yang tertua dan berguna sebagai sumber rujukan terkini (terutama bidang matematik). Di samping itu, teks klasik merupakan satu nilai ilmu yang mendidik dan menjelaskan moral terhadap kehidupan beragama (Jelani, 2004). Oleh hal tersebut, operasi teks klasik ini sebagai data yang mengandungi leksikal sebagai inventori

keilmuan masyarakat terdahulu (Chen, 2018). Pendedahan dibuat oleh sarjana matematik (Shaharir & Affandi, 2019) telah menjelaskan wujud leksikal per nomboran tertua hasil. Tambahan, struktur leksikal terdahulu yang menggunakan bahasa klasik telah membuktikan manuskrip Melayu beroperasi berlandaskan tulisan Jawi (Yakob, 2018). Proses transliterasi Jawi kepada rumi memberi manfaat terhadap pembangunan data raya korpus linguistik (Normi, 2020).

Berdasarkan daripada sumber maklumat teks klasik atau manuskrip ini, kajian ini mengkonsepkan teks klasik sebagai ilmu masyarakat Melayu feudal. Ilmu ini dirakam dengan penulisan tulisan Jawi dan telah dirumikan bagi kemudahan janaan dalam instrumen kajian. Teks klasik juga dikonsepskan sebagai himpunan bertulis yang merangkumi leksikal dan metafora bersifat kompleks dan memiliki catatan masyarakat terdahulu (Melayu). Teks klasik yang dihimpunkan dalam satu data korpus khusus iaitu Korpus Petua Membina Rumah (selepas ini disebut sebagai KPMR) meliputi 7 buah teks klasik Melayu (rujuk seksyen 3.2.3).

1.7.2 Ketumpatan Leksikal

Ketumpatan leksikal merupakan istilah dalam kaedah statistik pengukuran terhadap taburan kata sesebuah teks. Ketumpatan leksikal adalah kaedah untuk mengira bahagian perkataan leksikal dalam teks atau korpus dengan menyatakannya sebagai peratusan (Baker, Hardie, & McEnery, 2006). Ketumpatan leksikal ini boleh dilihat



menerusi Type Token Ratio (TTR). Namun, terdapat juga pengkaji melihat kekerapan mentah dan relatif (normalisasi) sebagai ketumpatan leksikal seperti kajian (Baayen, Milin, & Ramscar, 2016; Backman, 1978; Bradac, Desmond, & Murdock, 1977; Harald Baayen & Lieber, 1996; Šišková, 2012). Menurut Gregori-Signes & Clavel-Arroitia (2015) pula, fungsi ketumpatan leksikal atau *lexical density* merupakan satu istilah yang digunakan bagi mengukur statistik leksikal yang kaya dalam sesebuah teks. Kaedah ini menggunakan penjanaan perisian berdasarkan bidang komputer linguistik atau korpus linguistik. Berdasarkan ketumpatan leksikal ini mengandungi kekerapan leksikal atau *lexical frequency* yang memaparkan statistik tekstual leksikal (Arnold, Ballier, Lissón, & Tilton, 2019). Dalam kerencaman leksikal, boleh diperincikan wujud leksikal literal dan leksikal metafora yang merupakan senario normal tekstual (Feifei, 2018). Leksikal metafora merupakan satu bentuk kata yang mengandungi dua makna (rujuk seksyen 1.7.3), iaitu makna yang berlapis (harfiah dan makna lain). Oleh hal demikian, ketumpatan leksikal merupakan satu istilah yang memerihalkan statistik leksikal yang mengandungi leksikal literal/harfiah dan leksikal metafora.

1.7.2.1 Analisis Berpandukan Korpus

Analisis berpandukan korpus atau *corpus-driven analysis* merupakan satu pendekatan statistik leksikal, di mana korpus berdiri sendiri bersama-sama perisian dalam menganalisis data bahasa tabii. Menurut Biber et al. (2006:8), ahli bahasa berpandukan korpus dikatakan komited untuk 'integriti data sebagai keseluruhan' dan oleh itu dalam pendekatan yang terakhir ini, ia mendakwa bahawa kenyataan teori sepenuhnya





konsisten dengan mencerminkan secara langsung bukti yang disediakan oleh korpus. Dalam analisis ini terdapat tiga perkara utama, iaitu analisis kekerapan kata, analisis kata kunci dan analisis penghubungan matriks.

Analisis kekerapan kata adalah satu bentuk senarai kata yang digunakan dalam melihat kewujudan leksikal yang terkandung dalam teks. Analisis ini memiliki dua kekerapan, iaitu kekerapan mentah dan kekerapan relatif (Biber et al., 2006). Kekerapan ini merupakan panduan dalam menentukan perbandingan leksikal yang terdapat dalam teks (Crawford & Csomay, 2016). Kekerapan mentah merupakan nilai kiraan secara mentah berpandukan bilangan perkataan dalam sesebuah teks. Manakala kekerapan relatif merupakan nilai kiraan yang telah dinormalisasikan bergantung kepada per perkataan seperti 1000, 10,000 dan nilai gandaan mengikut kapasiti normalisasi korpus (Brezina, 2018). Hasil analisis ini dapat mengemuka tiga bentuk senarai kekerapan kata, iaitu kekerapan tertinggi, sederhana dan rendah (*hapaxes*).

Analisis kata kunci atau *keywords analysis* adalah penelitian perbezaan antara dua korpus iaitu korpus khusus dan korpus rujukan. Kata kunci didefinisikan sebagai satu leksikal yang berlaku kekerapan lebih tinggi (nilai signifikan) dalam satu korpus khusus berbanding korpus rujukan (O'Keeffe & McCarthy, 2012). Penjanaan ini merupakan satu analisis perbandingan dalam senarai kata korpus khusus (korpus kajian) dengan kekerapan korpus yang dibandingkan (Imran Ho-Abdullah, 2011). Janaan ini akan memerikan nilai keutamaan (*keyness*) setiap perkataan melalui pengujian statistik *Simple Maths Procedure (SMP)*.



Analisis penghubungan matriks (*co-occurrence matrik analysis*) merupakan satu representatif graf visualisasi yang menekankan plot leksikal terhadap kolokasi teks korpus (Brezina, 2018). Analisis ini menggunakan jaringan kolokasi bagi menghuraikan kedudukan leksikal dan plot leksikal berdasarkan senario kolokasi leksikal dalam data. Hasil analisis ini memberikan janaan ketumpatan leksikal berdasarkan skor yang telah ditetapkan. Analisis ini mampu menghasilkan interpretasi data kekerapan kolokasi secara geomatriks visual di mana setiap leksikal dipetakan dalam satu graf jaringan representatif (Brezina, 2018). Penilaian yang digunakan ialah logDice dengan nilai signifikan < 8.0 .

Dalam korpus linguistik terdapat istilah *representativeness* ialah salah satu tuntutan utama yang boleh dibuat untuk membina korpus bagi memperlihatkan contoh leksikal dari pelbagai bahasa tertentu (Baker et al., 2006). Pelbagai kaedah boleh dijelaskan untuk memerihalkan representasi antaranya sampel yang digunakan untuk memastikan tahap maksimum leksikal dalam reka bentuk pembinaan korpus (Coulthard, 2013). Melalui perwakilan atau representasi leksikal yang diinterpretasikan dapat memberi penjelasan terhadap kandungan korpus. Menerusi kajian ini, istilah representasikan digunakan bagi menjelaskan gambaran perwakilan leksikal yang mempunyai hubungan kolokasi melalui plot graf geomatriks/kolokasi (Brezina, 2018). Setiap leksikal yang dijelaskan mewakili data yang besar kerana leksikal tersebut telah disenaraikan dalam



kekerapan kata. Pengulangan/kekerapan ini tidak memberi kesan terhadap makna leksikal sebaliknya perubahan makna berlaku apabila terdapat hubungan kolokasi.

1.7.3 Metafora

Berdasarkan definisi dan ulasan Aristotle dalam karya Poetica (dalam tempoh 2300 tahun terdahulu) menyatakan “metafora terbentuk apabila sesuatu benda diberi nama (makna) yang dimiliki oleh benda lain; pemindahan (makna) ini berlaku sama ada daripada peringkat genus kepada jenisnya, atau daripada jenis kepada genusnya atau daripada jenis kepada jenisnya, atau berdasarkan analogi” (Hassan, 2016). Dalam pada itu, Cruse (2002) berpandangan bahawa metafora merupakan makna literal yang menggantikan makna literal pada perkataan asal, melalui metafora manusia dalam melihat sesuatu berdasarkan sesuatu yang lain. Selain itu, metafora menyediakan perspektif di dunia yang membantu membinanya dengan cara tertentu, menekankan ciri-ciri utama tertentu dan mengabaikan orang lain (Andriessen, 2007). Dalam kajian ini, metafora yang diketengahkan adalah leksikal metafora yang memiliki daya dalam menyampaikan lebih daripada dua makna dan hipotesis dalam representatif (Feifei, 2018). Dengan itu, metafora merupakan satu konsep leksikal yang padat dengan pelbagai makna, seterusnya menjadikan leksikal sebagai aras percambahan makna berlapis.





1.7.3.1 Prosedur Pengenalpastian Metafora

Prosedur Pengenalpastian Metafora atau lebih dikenali sebagai *Metaphor Identification Procedure (MIP)* merupakan satu tatacara dalam mengenal pasti metafora dalam sesebuah teks (Pragglejaz, 2007). Prosedur ini memiliki empat langkah dalam penentuan leksikal unit yang mewakili metafora (Semino, 2008). Bermula daripada langkah pembacaan sehingga langkah penentuan metafora, semua aspek diteliti terutama konteks leksikal sebagai penanda metafora.

Leksikogramartikal atau *lexicogram* merupakan satu pola metafora yang ditandai oleh variasi leksikal yang berkolokasi. Pola metafora ini melibatkan kemunculan leksikal yang mempengaruhi struktur metafora yang terkandung dalam teks. Pola metafora ini boleh ditentukan menerusi peringkat leksikal atau frasa (Sardinha, 2011). Di samping itu, terdapat juga sarjana memilih untuk meneliti leksikogramatikal ini daripada perspektif kolokasi (Fontaine, 2017). Di mana leksikal diuji dengan leksikal yang bersebelahan kiri dan kanan berdasarkan *Keywords In Context (KWIC)*. Hasil daripada itu, representatif ketumpatan leksikal dalam leksikogramatikal dikenal pasti dipetakan polanya.

Kebarangkalian metafora atau *metaphor probabilities* adalah satu nilai representatif unit leksikal yang mendorong kepada perkataan yang digunakan sebagai metafora atau *metaphor used word (MUW)* (Sardinha, 2011). Kiraan kebarangkalian metafora ini menunjukkan nisbah ketumpatan representatif leksikal terhadap



pembentukan metafora. Selepas itu, penentuan kebarangkalian perkataan yang digunakan sebagai metafora dalam leksikal yang berbeza dalam korpus yang sama. Kemudiannya, nilai representatif secara bandingan kemunculan leksikal digunakan sebagai metafora dalam korpus rujukan.

Dimensi variasi metafora adalah kepelbagaian metafora yang terkandung dalam sesebuah teks. Metafora ini dikenal pasti dengan menggunakan ketumpatan metafora merupakan kesinambungan daripada kebarangkalian metafora (Sardinha, 2011). Pada peringkat ini, huraian konteks digunakan berpandukan konkordans teks. Ketumpatan metafora dipilih berdasarkan pemboleh ubah metafora dan linguistik dalam literatur terdahulu seperti kajian Adams (2017); Rutten, Krogull, & Schoemaker, (2019); dan Sardinha (2011).

1.7.4 Ilmu Seni Bina Melayu

Ilmu seni bina Melayu merupakan disiplin mengenai industri pembinaan yang dibahagikan kepada sains bangunan dan teknologi binaan bangunan. Seni bina mempunyai konsep yang sangat luas dan boleh didefinisikan mengikut ruang lingkungan kajian. Tajuddin (1999), menyatakan seni bina dianggap sesuatu bentuk budaya seni dan kraf yang halus dan teliti yang mempunyai reka bentuk, corak tertentu, teknik pembinaan berkisarkan pembinaan bangunan-bangunan seperti masjid, istana, rumah bangsawan, serta makam diraja. Manakala Zulkifli Hanafi (1988), berpendapat

seni bina merangkumi berbagai-bagai bentuk dan ruang yang wujud untuk kegunaan manusia. Utaberta, Othman, & M. Surat, (2010) menyatakan pandangan bahawa alam bina memberi erti sebagai satu keindahan yang dicetuskan melalui kemahiran dalam binaan, tulisan, ukiran dan pertukangan. Maka, melalui leksikal yang dikaji dapat memberikan satu kejelasan ilmu seni bina Melayu terdahulu. Leksikal ini merakamkan gabungan konteks ilmu seni bina Melayu berdasarkan catatan manuskrip.

1.8 Kesimpulan

Dalam bab ini kupasan ilmu manuskrip, metafora dan korpus linguistik diteliti dengan berpada. Kupasan ini membincangkan soal latar belakang yang menyokong kajian ketumpatan leksikal dalam manuskrip Melayu dilakukan. Perbincangan isu dan permasalahan kajian diketengahkan bagi meneliti sebarang kelompangan dan kewajaran dalam kajian ini. Kesenambungan itu, objektif dan soalan kajian disusun mengikut keperluan permasalahan kajian seterusnya disokong dengan kepentingan kajian ini menjurus kepada pencetusan pola intelektual teks manuskrip Melayu. Dalam memberi kefahaman kajian ini, definisi operasional dinyatakan supaya tidak canggung dalam memahami landasan dan penerokaan kajian.