

Hakcipta © tesis ini adalah milik pengarang dan/atau pemilik hakcipta lain. Salinan boleh dimuat turun untuk kegunaan penyelidikan bukan komersil ataupun pembelajaran individu tanpa kebenaran terlebih dahulu ataupun caj. Tesis ini tidak boleh dihasilkan semula ataupun dipetik secara menyeluruh tanpa memperolehi kebenaran bertulis daripada pemilik hakcipta. Kandungannya tidak boleh diubah dalam format lain tanpa kebenaran rasmi pemilik hakcipta.



**VISUALISASI POHON SINTAKSIS BERASASKAN MODEL DAN
ALGORITMA SINTAKS AYAT BAHASA MELAYU**



YUSNITA BINTI MUHAMAD NOOR

UUM
Universiti Utara Malaysia

**DOKTOR FALSAFAH
UNIVERSITI UTARA MALAYSIA
2018**



Awang Had Salleh
Graduate School
of Arts And Sciences

Universiti Utara Malaysia

PERAKUAN KERJA TESIS / DISERTASI
(Certification of thesis / dissertation)

Kami, yang bertandatangan, memperakukan bahawa
(We, the undersigned, certify that)

YUSNITA MUHAMAD NOOR

calon untuk Ijazah **PhD**
(candidate for the degree of)

telah mengemukakan tesis / disertasi yang bertajuk:
(has presented his/her thesis / dissertation of the following title):

**"VISUALISASI POHON SINTAKSIS BERASASKAN MODEL DAN ALGORITMA SINTAKS AYAT
BAHASA MELAYU"**

seperti yang tercatat di muka surat tajuk dan kulit tesis / disertasi.
(as it appears on the title page and front cover of the thesis / dissertation).

Bahawa tesis/disertasi tersebut boleh diterima dari segi bentuk serta kandungan dan meliputi bidang ilmu dengan memuaskan, sebagaimana yang ditunjukkan oleh calon dalam ujian lisan yang diadakan pada : **23 Mei 2016.**

*That the said thesis/dissertation is acceptable in form and content and displays a satisfactory knowledge of the field of study as demonstrated by the candidate through an oral examination held on:
May 23, 2016.*

Pengerusi Viva:
(Chairman for VIVA)

Prof. Dr. Suhaidi Hassan

Tandatangan
(Signature)

Pemeriksa Luar:
(External Examiner)

Prof. Dr. Ahmad Zaki Abu Bakar

Tandatangan
(Signature)

Pemeriksa Dalam:
(Internal Examiner)

Dr. Siti Sakira Kamaruddin

Tandatangan
(Signature)

Nama Penyelia/Penyelia-penyelia: Prof. Dr. Zulikha Jamaludin
(Name of Supervisor/Supervisors)

Tandatangan
(Signature)

Nama Penyelia/Penyelia-penyelia: Dr. Shaidah Jusoh
(Name of Supervisor/Supervisors)

Tandatangan
(Signature)

Tarikh:
(Date) **May 23, 2016**

Kebenaran Mengguna

Penyerahan tesis ini, bagi memenuhi syarat sepenuhnya untuk ijazah lanjutan Universiti Utara Malaysia, saya bersetuju bahawa perpustakaan universiti boleh secara bebas membenarkan sesiapa sahaja untuk memeriksa. Saya juga bersetuju bahawa penyelia saya atau ketiadaannya, Dekan Awang Had Salleh Graduate School of Arts and Sciences diberi kebenaran untuk membuat salinan tesis ini dalam sebarang bentuk sama ada keseluruhannya atau sebahagiannya, bagi tujuan kesarjanaan. Adalah tidak dibenarkan sebarang penyalinan atau penerbitan atau kegunaan tesis ini sama ada sepenuhnya atau sebahagiannya bagi tujuan keuntungan kewangan/komersial, kecuali setelah mendapat kebenaran bertulis. Juga dimaklumkan bahawa pengiktirafan harus diberikan kepada saya dan Universiti Utara Malaysia dalam sebarang kegunaan kesarjanaan terhadap sebarang petikan daripada tesis saya.

Sebarang permohonan untuk menyalin atau menggunakan mana-mana bahan dalam tesis ini, sama ada sepenuhnya atau sebahagiannya hendaklah dialamatkan kepada:

Dekan Awang Had Salleh Graduate School of Arts and Sciences

UUM College of Arts and Sciences

Universiti Utara Malaysia

06010 UUM Sintok

Abstrak

Kajian terdahulu yang menghasilkan output pohon sintaksis dikaji dan didapati tidak bercambah untuk membuat paparan output yang lain. Oleh itu, kajian ini bermatlamat untuk menghasilkan satu algoritma untuk peningkatan output pohon sintaksis yang mana komponen output tambahan berkaitan dapat dihasilkan. Komponen tambahan iaitu semakan ayat, cadangan pembedahan ayat, visualisasi pohon sintaksis (VPS), dan atribut perkataan. Kesemua komponen ini terlebih dahulu dimodelkan dalam satu pakej sebelum diterjemahkan kepada prototaip. Dari segi penggunaan rumus binaan ayat, pengkaji Bahasa Melayu (BM) sebelum ini telah menggunakan Rumus Struktur Frasa (RSF). Namun, RSF telah didapati sebagai rumus yang tidak universal. Oleh itu, penggunaan rumus *X-bar* dalam kajian VPS ayat BM menjadi antara sumbangan kajian ini. Untuk mencapai objektif kajian (algoritma, model dan rumus *X-bar*), terdapat lima fasa kaedah penyelidikan terlibat. Fasa ini meliputi fasa pengetahuan pernyataan masalah, fasa mengkategorikan dan menganalisis rumus binaan ayat, reka bentuk model dan algoritma, fasa pembangunan prototaip, dan fasa penilaian dan rumusan. Kaedah penilaian Parseval, yang merupakan kaedah penilaian output dalam pemprosesan bahasa semula jadi telah digunakan untuk penilaian. Titik analisa kajian adalah metrik penilaian *recall* dan *precision*. Hasil output VPS diperoleh dengan purata 100% bagi *recall* dan 97.8% *precision*. Manakala hasil output cadangan pembedahan ayat pula memperolehi 100% *recall* dan 87.8% *precision*. Hasil output ini membuktikan bahawa algoritma dan model output tambahan boleh dimanfaatkan untuk digunakan dalam bahasa yang lain. Penilaian pengguna juga turut dilakukan dengan peratusan kepuasan subjektif 87.9% dan skor min sebanyak 6.157 mengikut skala perbezaan semantik 1 hingga 7. Penilaian kognitif pula mencatat 84.6% dengan skor min 4.230 mengikut skala Likert 1 hingga 5. Hasil analisis ini menunjukkan skor positif diperolehi untuk produk berasaskan model terutama dari segi kebergunaan, kemudahan penggunaan, kemudahan pembelajaran, kepuasan subjektif dan kognitif. Oleh itu, dapat disimpulkan bahawa algoritma dan model yang dicadangkan adalah berguna untuk pembangunan prototaip. Prototaip tersebut boleh dijadikan sebagai bantuan pembelajaran dalam memahami pembentukan ayat BM apabila dibekalkan dengan output yang dipertingkatkan pada semakan ayat, cadangan pembedahan ayat, VPS dan atribut perkataan.

Kata kunci: Pengkomputeran linguistik, Pohon sintaksis, Visualisasi pohon hurai, Penghurai ayat Bahasa Melayu

Abstract

Previous works that produce syntactic tree output has disregarded additional relevant components such as sentence checking, sentence correction, the syntax tree visualization and the words attributes of each sentence. As such, this study aims at producing an algorithm for syntactic tree output enhancement from which the relevant output component mentioned above can be produced. The additional components namely sentence checking, sentence correction, syntax tree visualization (VPS) and word attribute are modelled into a package prior to translating them into a tangible output. In term of rules, previous studies have used phrase-structure rules (RSF) in analysing the Malay sentence. But RSF has been found to be a non-universal formula. Our work has brought us to the introduction of X-bar rules for BM VPS, which consequently becomes one of the contributions of this study. To achieve these objectives (the algorithm, the model and the X-bar rules), five phases of research methods involved namely identifying the research gap, the sentence and rules categorization, model and algorithm design phase, prototype development evaluation and conclusion phase. Parseval assessment method, which is an output evaluation method in natural language processing, was used for the evaluation. Point of analysis were the recall and precision valuation metrics. For VPS output, the average results obtained were 100% for recall and 97.8% for precision. For sentence correction, the results given were 100% for recall and 87.8% for precision. These results proved that the algorithm and model, for syntactic tree output enhancement, are generalisable enough to be tested on other languages. User evaluation on the prototype was also performed yielding in the average subjective satisfaction of 87.9% and a mean score of 6.157, based on semantic differential scales of 1 to 7. Cognitive assessment was also recorded, obtaining average cognitive score of 84.6% with a mean score of 4.230, on the scale 5. Analysis on those results indicated positive scores on the model-based product specifically on usefulness, ease of use, ease of learning, subjective satisfaction, and cognitive measures. It can be concluded that the algorithm and model proposed were useful for the development of the prototype. The prototype is therefore beneficial as an educational assistance to understand Malay sentences when provided with enhanced output on sentence checking, sentence correction, syntax tree visualization (VPS) and words attribute.

Keywords: Computational linguistic, Syntactic parser, Parse tree visualization, Malay sentence parsing

Penghargaan

Alhamdulillah syukur kepada Allah S.W.T. dengan izin-Nya saya berjaya menyelesaikan kajian ini.

Setinggi-tinggi penghargaan diucapkan kepada penyelia saya Prof. Dr. Zulikha binti Jamaludin yang banyak bersabar dan tidak pernah jemu memberi tunjuk ajar. Saya sangat bersyukur kerana mendapat penyelia seperti beliau kerana bukan hanya dapat berguru dengan seorang yang sangat berilmu dalam pelbagai bidang, malah kualiti ilmu yang diperoleh juga membuka minda saya untuk menjadi orang yang lebih baik. Beliau seorang penyelia yang terbaik dalam kalangan yang terbaik.

Setinggi penghargaan juga kepada Munsyi Dewan Puan Ros Silawati binti Ahmad, Puan Siti Salmah binti Sulaiman, Puan Noor Suraya binti Adnan Sallehudin dan Prof Madya Dr. Wan Amizah binti Wan Mahmud yang terlibat dalam pengesahan rumus, ayat dan model kajian ini. Juga kepada Dr. Sabrina Tiun dan Prof Madya Dr. Nazlia Omar dari Universiti Kebangsaan Malaysia yang turut terlibat dalam membuat penentusahan model VPS dengan output tambahan.

Jutaan terima kasih juga kepada Dr. Nazihah binti Ahmad dari Pusat Pengajian Sains Kuantitatif UUM di atas bantuan dalam menghasilkan algoritma berbentuk persamaan matematik dalam kajian ini. Juga kepada Encik Alkaha bin Romli yang banyak membantu dalam proses pembangunan prototaip kajian.

Ucapan jutaan terima kasih kepada Awang Had Salleh UUM CAS atas setiap bantuan yang diberikan. Juga kepada Kementerian Pendidikan Tinggi Malaysia (MyPhd) yang menyediakan biasiswa pengajian saya (2012-2013), terima kasih diucapkan.

Suami Mansur bin Ismail dan anak-anak (Maisarah, Arsyad, Yariqa), serta semua sahabat, terima kasih semua.

Senarai Kandungan

Kebenaran Mengguna	ii
Abstrak	iii
<i>Abstract</i>	iv
Penghargaan	v
Senarai Kandungan	vi
Senarai Jadual.....	xi
Senarai Rajah	xiii
Glosari.....	xvi
Senarai Singkatan.....	xviii
BAB SATU PENGENALAN KAJIAN.....	1
1.0 Pengenalan.....	1
1.1 Pernyataan Masalah.....	2
1.2 Persoalan Kajian	5
1.3 Objektif Kajian	6
1.4 Skop Kajian	8
1.4.1 Skop Domain	8
1.4.2 Skop Struktur Ayat	9
1.4.3 Skop VPS.....	10
1.5 Reka Bentuk Kajian.....	12
1.6 Sumbangan Kajian.....	13
1.6.1 Model VPS dengan Output Tambahan	14
1.6.2 Algoritma VPS dengan Semakan serta Cadangan Pembetulan Ayat ...	14
1.6.3 Rumus Binaan Ayat	15
1.6.4 Pengecaman Atribut Perkataan.....	15
1.7 Struktur Tesis.....	16
1.8 Rumusan Bab Satu.....	18
BAB DUA ULASAN KARYA.....	20
2.0 Pengenalan.....	20

2.1	Latar Belakang Kajian	21
2.1.1	Kajian Pemprosesan Ayat di Malaysia	23
2.1.2	Penghurai Sintaksis	24
2.2	Kerangka Teori	25
2.2.1	Teori Graf	27
2.2.2	Teori <i>X-bar</i>	32
2.2.3	Teori Gestalt	33
2.2.4	Teori Beban Kognitif	34
2.3	Kajian Berkaitan Rumus Binaan Ayat BM	36
2.4	Kajian Berkaitan Model, Algoritma dan Prototaip	38
2.4.1	Model	41
2.4.1.1	Model SSTC	41
2.4.1.2	Model Penghurai Ayat	43
2.4.1.3	Model Penghurai Semantik	43
2.4.1.4	Implikasi daripada Model Terdahulu	44
2.4.2	Algoritma Penghurai Ayat	45
2.4.2.1	Penghurai Lehner's Prolog Tree Drawing	45
2.4.2.2	Penghurai phpSintakTree	46
2.4.2.3	Penghurai SynView	47
2.4.2.4	Penghurai RSyntaxTree	48
2.4.2.5	Penghurai Ayat Bahasa Melayu	49
2.4.2.6	Penghurai Ayat Bahasa Arab	52
2.4.2.7	Penghurai Statistik Ayat Bahasa Melayu	53
2.4.2.8	Penghurai Statistik Ayat Bahasa Myanmar	54
2.4.2.9	Penghurai Statistik Ayat Korea	55
2.4.2.10	Penghurai Tatabahasa Link Grammar	56
2.4.2.11	Implikasi daripada Penghurai Ayat	57
2.4.3	Penyemak Ayat	59
2.4.3.1	Penyemak Sintak Bahasa Melayu	60
2.4.3.2	Penyemak Ayat Bahasa Melayu	60
2.4.3.3	Penyemak Sintak Ayat BI	62

2.4.3.4	Implikasi daripada Penyemak Ayat.....	62
2.4.4	Visualisasi Struktur Ayat: Ekstrak Visualisasi Teks Subjektif.....	64
2.4.5	Cadangan Pembetulan Ayat.....	65
2.4.6	Atribut Perkataan: MALEX.....	65
2.4.7	Implikasi Kajian Berkaitan.....	66
2.5	Penentusahan Model dan Pembuktian Konsep.....	70
2.5.1	Penentusahan Model Kajian Terdahulu.....	70
2.5.2	Pembuktian Konsep Kajian Terdahulu.....	71
2.6	Jurang Kajian.....	73
2.7	Rumusan Bab Dua.....	79
BAB TIGA METODOLOGI KAJIAN.....		81
3.0	Pengenalan.....	81
3.1	Fasa 1: Pengetahuan Pernyataan Masalah.....	86
3.1.1	Kajian Awalan.....	86
3.1.2	Kajian Karya Terdahulu dan Analisis Kandungan.....	87
3.1.3	Kajian Perbandingan Karya Terdahulu.....	87
3.2	Fasa 2: Rumus <i>X-bar</i> , Model dan Algoritma.....	88
3.2.1	Pengumpulan Ayat.....	89
3.2.2	Pengesahan Lakaran.....	90
3.2.3	Kumpul Atribut Perkataan.....	92
3.2.4	Reka Bentuk Pangkalan Data.....	93
3.2.5	Reka Bentuk Model dan Algoritma.....	94
3.3	Fasa 3: Pembangunan Prototaip.....	95
3.4	Fasa 4: Penilaian Prototaip.....	98
3.4.1	Kaedah Penilaian Parseval.....	100
3.4.2	Kaedah Penilaian Pengguna.....	103
3.5	Rumusan Bab Tiga.....	107
BAB EMPAT MODEL DAN ALGORITMA.....		109
4.0	Pengenalan.....	109
4.1	Pembangunan Model.....	110
4.1.1	Model Atribut Perkataan.....	110

4.1.2	Model VPS dengan Output Tambahan	111
4.1.3	Penentuan Model VPS dengan Output Tambahan	118
4.2	Pembangunan Algoritma	124
4.2.1	Algoritma VPS dengan Output Tambahan	125
4.3	Rumusan Bab Empat	130
BAB LIMA PEMBANGUNAN, LATIHAN, PENILAIAN PROTOTAIP DAN PERBINCANGAN.....		132
5.0	Pengenalan.....	132
5.1	Aplikasi Teori Gestalt dan Teori Beban Kognitif Dalam Prototaip	132
5.2	Reka Bentuk dan Pembangunan Prototaip	134
5.2.1	Menghubungkan Pangkalan Data dan Antara Muka	136
5.2.2	Pengekoden Atur Cara Pembangunan VPS	141
5.3	Antara Muka Prototaip	143
5.3.1	Token Perkataan dan Semak Bilangan Perkataan.....	143
5.3.2	Semak Syarat Ayat, Penandaan Kelas Kata, Semak Ejaan.....	144
5.3.3	Semak Rumus, Cadangan, VPS	145
5.3.4	Atribut Perkataan	147
5.4	Output Tambahan	148
5.4.1	Rumus <i>X-bar</i>	148
5.4.2	Cadangan Pembetulan Ayat.....	149
5.4.3	Atribut Perkataan	149
5.4.4	VPS Ayat Contoh.....	149
5.5	Latihan Prototaip	150
5.6	Penilaian Prototaip.....	154
5.6.1	Penilaian Parseval	154
5.6.2	Ayat Uji Kaji.....	155
5.6.3	Hasil Uji Kaji VPS.....	155
5.6.4	Hasil Uji Kaji Cadangan Pembetulan Ayat	164
5.6.5	Penilaian Pengguna.....	167

5.7	Perbincangan Dapatan	176
5.8	Rumusan Bab Lima	183
BAB ENAM RUMUSAN.....		185
6.0	Pengenalan.....	185
6.1	Rumusan Sumbangan Kajian.....	185
6.2	Rumusan Pencapaian Objektif.....	187
6.3	Kekangan Kajian	190
6.4	Penambahbaikan Masa Hadapan	192
	Rujukan.....	194



UUM
Universiti Utara Malaysia

Senarai Jadual

Jadual 1.1 Reka Bentuk Kajian	13
Jadual 2.1 Kajian Berkaitan	40
Jadual 2.2 Ringkasan Sorotan Karya Tentang Pembangunan Model	44
Jadual 2.3 Ringkasan Sorotan Karya Tentang Algoritma atau Kaedah	58
Jadual 2.4 Ringkasan Sorotan Karya Tentang Penyemak Ayat.....	63
Jadual 2.5 Ringkasan Kajian Terdahulu.....	67
Jadual 3.1 Jumlah Ayat yang Diasingkan	90
Jadual 3.2 Pembahagian Ayat	92
Jadual 3.3 Kaedah dan Teknik Penilaian Reka Bentuk.....	98
Jadual 3.4 Metrik Penilaian Penghurai Ayat.....	102
Jadual 3.5 Hasil Kebolehpercayaan Instrumen Kajian	106
Jadual 4.1 Komponen Model VPS dengan Output Tambahan	114
Jadual 4.2 Peringkat Pemprosesan Model Piramid	117
Jadual 4.3 Hasil Penentusahan Model.....	120
Jadual 4.4 Hasil Penilaian Komponen	121
Jadual 4.5 Cadangan Penambahbaikan Model VPS dengan Output Tambahan.....	123
Jadual 5.1 Jumlah Ayat Untuk Latihan Prototaip	150
Jadual 5.2 Hasil Uji Kaji Fasa Latihan Prototaip Secara Keseluruhan	152
Jadual 5.3 Rumus <i>X-bar</i>	153
Jadual 5.4 Bilangan Ayat Uji Kaji Mengikut Pola Ayat BM.....	155
Jadual 5.5 Ringkasan Hasil Uji Kaji VPS.....	156
Jadual 5.6 Purata dan Peratus Hasil Uji Kaji VPS	156
Jadual 5.7 Ayat Output Melebihi Satu	157
Jadual 5.8 Contoh Ayat dengan Penggunaan Unsur Penerang dalam Subjek	159
Jadual 5.9 Ayat dengan Cadangan yang Salah.....	164
Jadual 5.10 Cadangan Pembetulan Ayat.....	165
Jadual 5.11 Ringkasan Hasil Uji Kaji Cadangan Pembetulan Ayat.....	166
Jadual 5.12 Purata dan Peratus Hasil Uji Kaji Cadangan Pembetulan Ayat.....	167
Jadual 5.13 Hasil Keseluruhan Berdasarkan Soal Selidik USE.....	168

Jadual 5.14 Soalan Bagi Skala Minimum 1 Soal Selidik USE	171
Jadual 5.15 Hasil Min Penilaian Soal Selidik USE Mengikut Tingkatan.....	172
Jadual 5.16 Hasil Penilaian Kognitif.....	173
Jadual 5.17 Hasil Min Penilaian Kognitif Mengikut Tingkatan	175



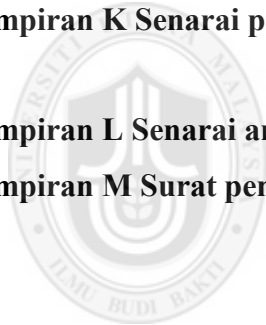
Senarai Rajah

Rajah 1.1. Skop kajian	11
Rajah 1.2. Struktur tesis berdasarkan objektif	16
Rajah 2.1. Struktur sains linguistik menunjukkan kaitan CL dan SL melalui LU.....	22
Rajah 2.2. Pendekatan graf berhierarki	28
Rajah 2.3. Perkaitan teori graf dengan skop kajian.....	31
Rajah 2.4. Kerangka teori.....	35
Rajah 2.5. Model penganalisis sintak SSTC	42
Rajah 2.6. Penghurai Lehner's Prolog Tree Drawing.....	46
Rajah 2.7. Penghurai phpSintakTree	47
Rajah 2.8. Penghurai SynView	48
Rajah 2.9. Penghurai RSyntaxTree	49
Rajah 2.10. Contoh output Penghurai ayat Bahasa Melayu.....	50
Rajah 2.11. Antara muka sistem penghurai ayat Bahasa Melayu	51
Rajah 2.12. Penghurai ayat bahasa Arab.....	52
Rajah 2.13. Output penghurai statistik bahasa Korea	56
Rajah 2.14. Output mengekstrak teks subjektif	64
Rajah 3.1. Perkaitan metodologi dengan sumbangan dan objektif kajian	83
Rajah 3.2. Metodologi kajian PR.	85
Rajah 3.3. Carta alir proses mengkategorikan dan menalisis ayat	89
Rajah 3.4. Perkataan disimpan dalam Fail Perkataan	94
Rajah 3.5. Carta alir proses menghasilkan model dan algoritma.....	94
Rajah 3.6. Seni bina prototaip VPS dengan output tambahan.	97
Rajah 4.1. Komponen model dan teori.....	110
Rajah 4.2. Model atribut perkataan	111
Rajah 4.3. Perkaitan komponen model VPS dengan output tambahan dan teori.....	112
Rajah 4.4. Model VPS dengan output tambahan (model piramid)	113
Rajah 4.5. Empat sisi model piramid	116
Rajah 4.6. Model VPS dengan output tambahan ditambah baik.....	124

Rajah 4.7. Perkaitan komponen model dan kaedah kajian.....	125
Rajah 4.8. Langkah algoritma VPS dengan output tambahan	126
Rajah 4.9. Carta alir VPS	128
Rajah 4.10. Carta alir atribut perkataan	129
Rajah 4.11. Carta alir VPS ayat contoh.....	130
Rajah 5.1. Proses menganalisis ayat	135
Rajah 5.2. Keratan fail rumus.cfg.	138
Rajah 5.3. Keratan fail perkataan.cfg.....	138
Rajah 5.4. Keratan fail imej	139
Rajah 5.5. Keratan fail ayat contoh.cfg.....	139
Rajah 5.6. Paparan senarai contoh ayat.....	140
Rajah 5.7. Keratan fail ayat majmuk.cfg.....	141
Rajah 5.8. Antara muka BMTutor.....	142
Rajah 5.9. Semak bilangan perkataan	143
Rajah 5.10. Semak syarat ayat	144
Rajah 5.11. Perkataan yang tiada dalam simpanan	144
Rajah 5.12. Ayat yang tidak dapat diproses.....	145
Rajah 5.13. Cadangan pembetulan ayat.....	145
Rajah 5.14. VPS ayat input	146
Rajah 5.15. Atribut perkataan dan VPS ayat contoh.....	147
Rajah 5.16. Contoh output frasa nama yang diasingkan.....	159
Rajah 5.17. Contoh VPS	160
Rajah 5.18. Contoh VPS	161
Rajah 5.19. Contoh VPS	162
Rajah 5.20. Contoh VPS	163

Senarai Lampiran

Lampiran A Aplikasi pohon sintaksis untuk BI	203
Lampiran B Contoh surat persetujuan responden	208
Lampiran C Surat persetujuan pengetua	209
Lampiran D Surat kebenaran pengumpulan data.....	210
Lampiran E Instrumen penilaian pakar.....	211
Lampiran F Carta alir VPS dengan output tambahan	214
Lampiran G Rumus <i>X-bar</i>	217
Lampiran H Ayat uji kaji.....	219
Lampiran I Hasil uji kaji cadangan pembetulan ayat.....	223
Lampiran J Biodata penilai pakar.....	245
Lampiran K Senarai penerbitan..	246
Lampiran L Senarai anugerah.....	247
Lampiran M Surat pengesahan Munsyi Dewan	248



UUM
Universiti Utara Malaysia

Glosari

Istilah yang sering digunakan dalam penulisan tesis ini adalah model, algoritma, visualisasi dan pohon sintaksis. Setiap istilah ini dijelaskan seperti berikut.

Model

Jurafsky dan Martim (2009) menyatakan bahawa model mengandungi komponen, perkaitan antara komponen dan persembahan. Manakala Hunter (2006) menyatakan bahawa model adalah gambaran konsep tentang komponen yang mempersembahkan pengetahuan dalam memahami proses penyambungan aliran data. Oleh itu, dalam kajian ini, model diertikan sebagai kombinasi komponen dan perkaitan antara komponen yang membentuk sebuah model bagi mereka bentuk prototaip.

Algoritma

Algoritma didefinisikan sebagai prosedur perkomputeran untuk mencapai perkaitan antara input dan output (Cormen, Leiserson, Rivest & Stein, 2001). Algoritma dijelaskan dalam bentuk turutan berbentuk kod pseudo atau carta alir pembangunan (Voloshin, 2009; Yuni Dwi, 2005). Oleh itu, dalam kajian ini, algoritma adalah turutan VPS dengan output tambahan yang diterjemahkan dalam bentuk persamaan matematik dan carta alir.

Pohon Sintaksis

Pohon sintaksis merupakan penanda frasa dan dianggap sebagai keterangan struktur pembentukan ayat (Nik Safiah, Farid, Hashim & Abdul Hamid, 2009).

Visualisasi

Visualisasi adalah konsep mempersembahkan aliran data dan pembangunan. Ia adalah salah satu kaedah yang dapat membantu menghuraikan data yang sukar (Ware, 2000; Kaidi, 2000; Bjork, Holmquist & Redstrom (n.d)). Dalam kajian ini, istilah visualisasi adalah merujuk kepada kaedah paparan pohon sintaksis yang dinamakan sebagai visualisasi pohon sintaksis (VPS). VPS digunakan untuk menghuraikan pembentukan ayat penyata BM.



Senarai Singkatan

BM	Bahasa Melayu
BI	Bahasa Inggeris
BMTutor	Bahasa Melayu Tutor
DBP	Dewan Bahasa dan Pustaka
N"	Frasa nama
SN"	Frasa nama subjek
K"	Frasa kerja
A"	Frasa adjektif
KS"	Frasa sendi nama
N	Kata nama
K	Kata kerja
A	Kata adjektif
KS	Kata sendi nama
KT	Kata tugas
N'	Frasa pertengahan N"
K'	Frasa pertengahan K"
A'	Frasa pertengahan A"
KS'	Frasa pertengahan KS"
PK	Penerang kata kerja
PA	Penerang kata adjektif
KBII	Kata bilangan
KB	Kata bantu
KAD	Kata adverba
KNF	Kata nafi
KPM	Kata pemerih
KP	Kata penguat
KPN	Kata penegas
LG	<i>Link Grammar</i>
PENT	Penentu

RSF	Rumus struktur frasa
CFG	<i>Context-free grammar</i>
SSTC	<i>Structure-String Tree Correspondence</i>
VPS	Visualisasi pohon sintaksis



BAB SATU

PENGENALAN KAJIAN

1.0 Pengenalan

Kajian dalam bidang pengkomputeran linguistik semakin berkembang di Malaysia. Hal ini telah menghasilkan banyak aplikasi sealiran seperti kamus Dewan Eja, MALEX iaitu sebuah pangkalan data tatabahasa Bahasa Melayu (BM), mesin terjemahan dan pelbagai jenis kamus elektronik. Namun dalam menghasilkan pemprosesan tahap ayat terutama BM sebagai bahasa utama di Malaysia masih perlu diberi penekanan (Siti Hajar, 2011). Bantuan teknologi diperlukan agar penguasaan yang baik boleh diterapkan dan diperkembangkan (Sekretariat Pusat Majlis Bahasa Melayu IPT Nusantara, 2013). Selain itu, penutur BM perlu mempunyai ilmu dan hasil ciptaan sendiri dan tidak senantiasa berharap akan ehsan pencipta teknologi bahasa lain serta mempunyai kemahiran sendiri (Jaafar, 2008; Abdullah, 2010). Ramai penyelidik yang merungkai keperluan aplikasi pemprosesan bagi BM seperti yang dinyatakan dalam Zuraidah (2010), Mohd Juzaidin (2007;2008), dan Nazri, Muhammad, Shamsinah, Norizillah dan Fatahiyah (2006) dalam kajian tentang pengkomputeran linguistik dan bahasa tabii di Malaysia.

Salah satu kaedah yang telah diperkenalkan oleh pengkaji untuk menggambarkan struktur ayat adalah dalam bentuk rajah berpokok atau lebih dikenali sebagai pohon penghurai atau pohon sintaksis. Pohon sintaksis berkomputer atau visualisasi pohon sintaksis (VPS) telah diperkenalkan di Malaysia bagi ayat BM. VPS yang dihasilkan perlu dikembangkan supaya boleh dirujuk dan digunakan dalam aplikasi yang lain. VPS

diperlukan sebagai bantuan kepada aplikasi pemrosesan bahasa yang lain seperti pemrosesan semantik. Hal ini dinyatakan dalam Mohd Juzaidin (2007) tentang keperluan teknik pemrosesan BM. Oleh yang demikian, kajian berkaitan VPS dikaji untuk melihat penambahbaikan yang boleh dilakukan ke atas jurang yang dikenal pasti seperti dalam bahagian pernyataan masalah seterusnya.

1.1 Pernyataan Masalah

Kajian pemrosesan BM telah dikaji sejak tahun 1980an. Bermula dengan kajian berkenaan morfologi sehingga kajian penghuraian ayat rancak dilakukan hingga kini. Contohnya kajian mengekstrak teks BM diperkenalkan pada tahun 2014. Walau bagaimanapun, kajian terdahulu tidak menyentuh pohon sintaksis secara mendalam sebagaimana yang dapat dilihat dalam Noor Hafhizah (2011), Suzaimah (2002), Rosmah (1995), Ahmad Izuddin *et al.* (2007), Al-Adhaileh dan Kong (1998), Murugesan dan Cassimatis (2006), Peters (2008), Sleator dan Temperley (1993), Rozana, Nurul Atiqah, Eliza Mazmee dan Saipunidzam (2011), dan Zuraidah (2010).

Kajian-kajian Noor Hafhizah (2011), Ahmad Izuddin *et al.* (2007), Suzaimah (2002) dan Rosmah (1995) sebagai contoh, membuat semakan ayat dan hasil output yang diberikan adalah pohon sintaksis berbentuk hierarki atau berbentuk separa. Output yang dihasilkan terhad kepada pohon sintaksis tersebut. Kajian-kajian ini boleh ditambahbaik dengan tambahan elemen sokongan untuk pemahaman pengguna. Di antara penambahan yang mungkin adalah 1) cadangan pembedahan ayat, 2) atribut perkataan, dan 3) VPS ayat contoh. Walau bagaimanapun, model dan algoritma yang mendasari output tambahan

tersebut masih belum diperkenalkan. Sebarang penambahan bentuk output pohon sintaksis memerlukan model yang utuh supaya aplikasi yang berkaitan dengannya dapat digeneralisasikan. Oleh yang demikian, permasalahan kajian ini adalah ketiadaan model dan algoritma untuk output tambahan pohon sintaksis.

Cadangan pembetulan ayat dalam proses semakan perlu disertakan dalam VPS. Hingga kini, penyemak ayat BM (Rozana *et al.*, 2011) adalah kajian yang memberi cadangan kepada pengguna apabila terdapat kesalahan pada ayat input. Cadangan berupa kelas kata dipaparkan apabila ayat input didapati tidak sepadan dengan rumus yang disediakan. Contohnya, ayat yang dimasukkan tanpa kata sendi nama (KS) dalam ayat berpola frasa sendi nama akan mengeluarkan ralat menyatakan ketiadaan KS dalam ayat tersebut. Walau bagaimanapun, cadangan pembetulan ayat tidak dilakukan. Oleh itu, cadangan pembetulan ayat adalah salah satu output tambahan yang perlu dimodelkan.

VPS yang boleh membuat atribut perkataan juga masih baru diperkenalkan. Hingga kini, MALEX (MALay LEXicon) sebagai pangkalan data BM (Zuraidah, 2010) menjadi sumber rujukan dalam pemprosesan tatabahasa BM. Namun, MALEX tidak dihasilkan daripada VPS. MALEX juga sebuah pangkalan data dan bukan sebuah aplikasi VPS. Kerana itu, VPS dengan output tambahan seperti atribut perkataan diperlukan agar dapat memberi lebih banyak manfaat kepada pengguna.

Bagi menghasilkan VPS yang tepat paparannya mengikut struktur binaan ayat yang betul, maka rumus binaan ayat perlu dikaji. Sehingga kini, Noor Hafhizah (2011) telah

menyediakan Rumus Struktur Frasa (RSF) dalam pembangunan prototaip kajian beliau. Namun, RSF yang dibangunkan tidak mengikut rumus binaan ayat terkini yang dihasilkan oleh Dewan Bahasa dan Pustaka (DBP). Selain tidak universal, RSF yang dibangunkan juga terhad kepada 147 rumus. Selain RSF, masih tidak terdapat rumus binaan ayat BM dikaji untuk kegunaan VPS. Justeru itu, kajian ini diperlukan bagi mereka bentuk rumus lengkap untuk pembangunan VPS ayat penyata BM.

Model pemprosesan ayat telah diusulkan dalam aplikasi *structured string-tree correspondence (SSTC)* (Al-Adhaileh & Kong, 1998), penghurai sintaksis (Murugesan & Cassimatis, 2006), dan penghurai semantik (Peters, 2008). Model ini walau bagaimanapun tidak meliputi semakan atribut perkataan, pemprosesan semakan semula ayat, pembedaan ayat atau mencadangkan ayat yang betul. Komponen semakan dengan cadangan pembedaan ayat ini penting dalam penghasilan VPS yang tepat dan boleh membantu pengguna untuk mengetahui pembentukan struktur ayat yang betul.

Algoritma pemprosesan ayat ditunjukkan dalam Noor Hafhizah (2011), Ahmad Izuddin *et al.* (2007) dan dalam kajian aplikasi *Link Grammar* (Sleator & Temperley, 1993). Namun, turutan yang diusulkan adalah turutan asas dalam pemprosesan bahasa tabii (NLP). Turutan untuk semakan ayat, pembedaan ayat, serta rujukan atribut perkataan masih belum mempunyai sebarang algoritma.

Dengan adanya model yang lasak dan algoritma yang jelas, penyelesaian berkenaan VPS dapat digeneralisasikan. VPS tidak akan hanya menghasilkan pohon sintaksis semata,

malah dapat digunakan untuk kegunaan bidang lain seperti terjemahan perkataan menggunakan VPS. Model yang diperkenalkan dapat digabung atau dikembangkan kepada model bidang sealiran, contohnya model VPS untuk mengkategorikan jenis ayat BM. Selain itu, algoritma yang jelas penting agar turutan pelaksanaan boleh diguna pakai dalam kajian VPS berkaitan.

Walaupun tanpa output tambahan, kajian seperti Noor Hafhizah (2011), Suzaimah (2002), Rosmah (1995), Ahmad Izuddin *et al.* (2007), Al-Adhaileh dan Kong (1998), Murugesan dan Cassimatis (2006), Peters (2008), Sleator dan Temperley (1993), Rozana *et al.* (2011) dan Zuraidah (2010) merupakan platform yang baik bagi kajian ini.

Untuk menyelesaikan permasalahan yang dikenal pasti, persoalan kajian dijelaskan dalam bahagian seterusnya.

1.2 Persoalan Kajian

Bagi mengusulkan satu model VPS dengan komponen tambahan, beberapa persoalan perlu dijawab. Persoalan ini berdasarkan kepada permasalahan ketiadaan model dan algoritma output tambahan pohon sintaksis seperti berikut.

1. Adakah rumus binaan ayat telah diperkenalkan untuk VPS?

- Adakah RSF boleh digunakan untuk VPS?
- Adakah cara yang lebih baik daripada kaedah mendapatkan RSF untuk mengkategorikan struktur ayat dan mendapatkan rumus binaan ayat bagi kegunaan VPS?

2. Bagaimana model, algoritma dan prototaip output tambahan pohon sintaksis boleh dibina?
 - Bolehkah atribut perkataan dimodelkan dalam VPS?
 - Adakah algoritma pohon sintaksis telah diperkenalkan? Bolehkah algoritma ini ditambah komponen semakan dengan cadangan dan atribut perkataan?
 - Bolehkah model dan algoritma tersebut diterjemahkan kepada prototaip sebagai cara pembuktian konsep?
3. Bagaimana cara untuk memastikan model dan algoritma yang dihasilkan adalah tepat?

Persoalan tersebut perlu dijawab untuk mencapai objektif berikut.

1.3 Objektif Kajian

Tujuan utama kajian ini adalah untuk menghasilkan model dan algoritma output tambahan pohon sintaksis. Berdasarkan model tersebut, dapat diterbitkan pula satu panduan dalam pembentukan VPS yang dapat membuat penambahbaikan dalam pohon sintaksis dengan memberi cadangan pembedahan ayat dan atribut perkataan serta membuat VPS melalui ayat contoh. Panduan ini diterjemahkan dalam bentuk model dan algoritma yang berasaskan kepada rumus binaan ayat nyata BM. Tiga sub-objektif berikut perlu dicapai bagi menyempurnakan objektif utama.

1. Mengkategorikan dan menganalisis struktur ayat BM untuk mendapatkan rumus yang tepat.

2. Membina model dan algoritma VPS dengan output tambahan serta prototaip kajian sebagai alat pembuktian konsep.
3. Menilai output VPS dan cadangan pembetulan ayat berdasarkan metrik penilaian dalam pemprosesan bahasa tabii untuk menguji ketepatan output dan membuat penilaian pengguna bagi mencapai kepuasan subjektif dan penerimaan kognitif pengguna.

Semua objektif tersebut bertujuan untuk menghasilkan VPS bagi struktur ayat BM dengan tambahan output dalam paparan pohon sintaksis. Objektif ini dianggap berjaya jika boleh membuktikan hipotesis seperti dalam bahagian seterusnya.

Hipotesis kajian ini adalah seperti berikut.

H_1 : Rumus binaan ayat yang diperolehi daripada ayat penyata BM boleh digunakan untuk membuat VPS (Sebaliknya rumus yang diperolehi daripada ayat penyata BM tidak boleh digunakan untuk menghasilkan VPS).

H_2 : Model tambahan output yang direka bentuk boleh digunakan untuk menghasilkan algoritma berkaitan untuk membangunkan prototaip (Sebaliknya, model yang dicadangkan tidak dapat digunakan untuk mereka bentuk algoritma bagi pembinaan prototaip kajian).

H_3 : Skor kepuasan subjektif dan penerimaan kognitif pengguna boleh diukur (Sebaliknya skor kepuasan subjektif dan kognitif pengguna tidak dapat diukur).

Pencarian jalan penyelesaian dalam mencapai kesemua objektif adalah berlandaskan kepada skop kajian seperti berikut.

1.4 Skop Kajian

Skop kajian dikelaskan mengikut domain, struktur ayat, dan VPS. Setiap kategori skop dihuraikan di bahagian 1.4.1 hingga 1.4.3.

1.4.1 Skop Domain

Kajian pengkomputeran linguistik menyumbang kepada salah satu komponen dalam bidang ilmu interaksi manusia-komputer (HCI) iaitu visualisasi. Visualisasi boleh digunakan untuk membantu pemahaman, sebagai contoh menggunakan kaedah pohon sintaksis, dapat membantu pemahaman pembentukan struktur ayat. Struktur ayat yang dipilih sebagai domain kajian ini adalah ayat BM. Hal ini kerana, hasil akhir kajian dapat digunakan oleh penutur BM khususnya pelajar sekolah. Data terkini Kementerian Pendidikan menunjukkan pelajar adalah lemah dalam menguasai gramatis ayat dan jenis frasa. Selain itu, mereka juga menghadapi kesukaran dalam tatabahasa BM (Zaharani & Nor Hashimah, 2012; Nor Hashimah, Junaini & Zaharani, 2010; Bagavathy, 2005). Pemilihan domain ini selari dengan dasar inovasi dalam pendidikan di Malaysia yang memfokuskan agar bahan pengajaran di sekolah menggunakan IT dan multimedia bagi membantu pelajar (Zaini *et al.*, 2012).

Ayat penyata dalam buku teks BM tingkatan satu hingga tingkatan lima dipilih sebagai data kajian. Pemilihan bahan bacaan ini difokuskan bagi sekolah menengah kerana walaupun kaedah binaan ayat dan tatabahasa diajar sejak di bangku sekolah rendah (Nik

Hassan Basri, 2009; Kementerian Pendidikan Malaysia, 2003; Nawi, 2003), namun, di sekolah menengah, pembelajaran tatabahasa dan binaan ayat lebih ditekankan (Abd. Aziz, 2000). Selain itu juga, terdapat majoriti pelajar sekolah menengah tidak menguasai tatabahasa sehingga mereka meninggalkan alam persekolahan. Mereka mengalami masalah dalam pembentukan ayat yang betul dan tidak dapat membezakan jenis tatabahasa dan kelas kata (Abdul Rashid, 2004; Nawi, 2003).

Jika dilihat dari sudut perkembangan teknologi, terdapat banyak aplikasi berkomputer telah dibangunkan terutamanya oleh para pengkaji Barat seperti penyemak ejaan, penterjemahan berkomputer, kamus atas talian dan penghurai istilah (Mohd Juzaidin, 2007). Oleh itu, BM juga perlu dikembangkan selaras dengan bahasa utama dunia yang lain. BM perlu menggunakan kaedah visualisasi sebagai bantuan dalam pemahaman pembentukan ayat. Visualisasi ayat ini dilakukan dengan membenarkan interaksi dilakukan antara pengguna dan sistem.

1.4.2 Skop Struktur Ayat

Ayat BM dikategorikan kepada empat jenis iaitu ayat penyata, ayat tanya, ayat perintah dan ayat seruan. Kajian ini diskopkan kepada ayat penyata. Ayat penyata dipilih mengikut Ahmad Izuddin *et al.* (2007) dan Noor Hafhizah (2011) sebagai data yang digunakan dalam mendapatkan rumus dan menjalankan uji kaji. Jika kajian berjaya mencapai objektif, maka ia boleh dikembangkan kepada ayat jenis lain. Jumlah bilangan perkataan bagi setiap ayat tidak melebihi 14 mengikut spesifikasi yang ditetapkan oleh Abdullah (2008) sebagai ayat mudah dan sederhana. Oleh itu, sebagai asas memahami pembentukan ayat, justifikasi ini diambil kira.

Terdapat enam fasa peringkat pengetahuan dalam memahami bahasa tabii iaitu peringkat fonologi, morfologi, sintaksis, semantik, pragmatik dan wacana (Noor Hafhizah, 2011). Sintak dan semantik adalah aspek yang saling berkait rapat. Kajian ini hanya memberi fokus kepada aspek sintaksis sahaja kerana aspek sintaksis merupakan aspek utama yang harus dipentingkan berbanding aspek lain (Zulkifley, 2012). Selain itu, aspek sintaksis juga adalah kajian tentang struktur pembentukan ayat berbanding aspek semantik yang lebih menekankan tentang preposisi atau makna (Siti Hajar, 2009; Nik Hassan Basri, 2009).

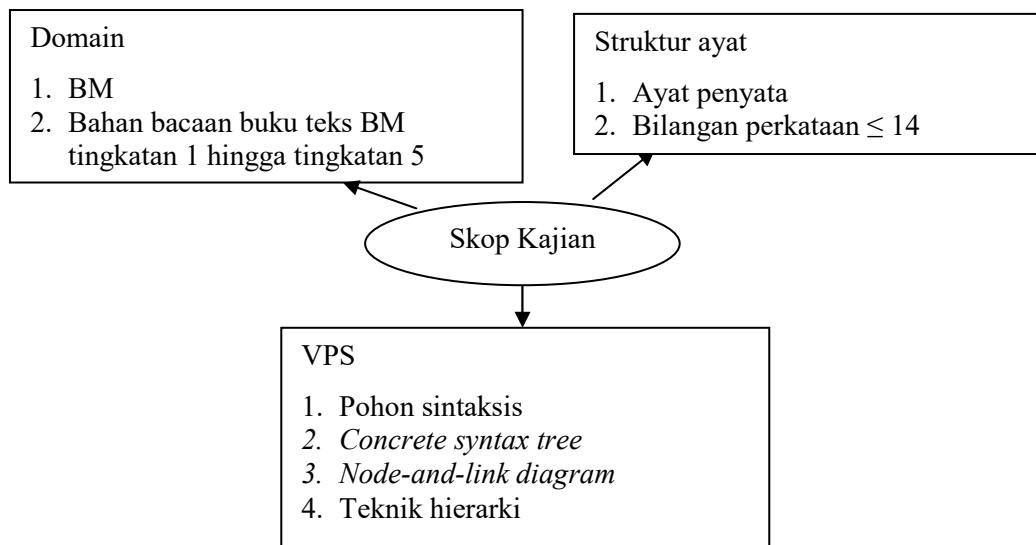
1.4.3 Skop VPS

Kaedah visualisasi dipilih kerana keberkesanannya dalam membantu pemahaman pelajar (Almeida-Martinez, Urquiza-Fuentes & Velzquez-Iturbide, 2009; Abdul Rahman Huraissen, 2012; Hamidah, 2010). Kaedah visualisasi ini boleh dilakukan menggunakan paparan pohon sintaksis. Pohon sintaksis terbahagi kepada dua jenis iaitu *Abstract syntax tree* (AST) atau lebih dikenali sebagai *syntax tree* yang dirujuk penggunaannya dalam bidang sains komputer dan *Concrete syntax tree* (CST) atau lebih dikenali sebagai pohon sintaksis yang dirujuk penggunaannya dalam bidang ilmu bahasa. AST tidak menunjukkan sintak secara terperinci, manakala CST adalah untuk ayat dalam bahasa tabii.

Kaedah menggunakan pohon sintaksis dipilih bagi tujuan memahami struktur binaan ayat. Penjelasan menggunakan pohon sintaksis adalah kaedah yang biasa digunakan oleh pengkaji bahasa seperti Nik Safiah *et al.* (2009), Hussin (n.d), Abdullah, Seri Lanang, Razali, dan Zulkifli (2006) dan Zaharin (1998). Kaedah ini pertama kali telah

diperkenalkan oleh Chomsky (1957). Selepas kaedah tersebut diperkenalkan, kebanyakan penerangan tentang ayat adalah dengan menggunakan pohon sintaksis.

Pohon sintaksis terbahagi kepada dua bentuk penghurai iaitu berbentuk struktur frasa (*phrase structure*) atau kebergantungan (*dependency*) (Kovar, 2014; Jakubicek, 2012). Penghurai secara kebergantungan menghurai ayat mengikut kebergantungan tatabahasa yang terlibat dalam ayat (Kakkonen, 2007). Manakala, penghurai berbentuk frasa digunakan untuk menghurai ayat secara hierarki (*deep parsing*) atau secara separa (*shallow/partial/chunking parsing*). Teknik hierarki akan menggunakan kaedah *node-and-link diagram* (Luboschik & Schumann, 2007; Phang & Zarina, 2012) bagi menghuraikan kedudukan setiap perkataan dalam ayat. Penghurai ini melibatkan penggunaan rumus binaan ayat sebagaimana ayat BM yang dibentuk menggunakan rumus. Oleh itu, penghurai berbentuk frasa adalah menjadi skop kajian ini. Rumusan tentang skop yang terlibat dalam kajian ini ditunjukkan dalam Rajah 1.1.



Rajah 1.1. Skop kajian

Skop kajian tersebut ditetapkan bagi membantu pengstrukturannya pelaksanaan kajian ini. Skop tersebut digunakan dalam fasa reka bentuk kajian seperti berikut untuk menjelaskan prosedur kajian terhadap domain, struktur ayat dan VPS.

1.5 Reka Bentuk Kajian

Terdapat lima fasa yang terlibat dalam rangka kerja kajian ini iaitu Fasa 1 menggunakan kaedah pengetahuan pernyataan masalah, Fasa 2 mencadangkan kaedah mengkategorikan ayat, rumus, reka bentuk model dan algoritma, Fasa 3 menggunakan kaedah pembangunan, Fasa 4 menjalankan kaedah penilaian dan Fasa 5 memberi rumusan. Fasa pengetahuan pernyataan masalah digunakan untuk menentukan domain kajian. Seterusnya, skop struktur ayat bagi mendapatkan rumus untuk kegunaan reka bentuk model dan algoritma dilakukan dalam Fasa kedua. Fasa pembangunan, penilaian dan rumusan adalah berdasarkan skop VPS yang difokuskan.

Setelah kajian difahami, ayat dan rumus dikategorikan untuk kegunaan fasa reka bentuk dan pembangunan. Fasa reka bentuk melibatkan aktiviti mereka bentuk model dan algoritma seperti model atribut perkataan dan model VPS dengan output tambahan. VPS tersebut perlu melalui proses penentusahan sebelum algoritma boleh direka bentuk. Seterusnya, prototaip BMTutor dibina pada fasa ketiga untuk pembuktian konsep dalam model. Prototaip diuji berdasarkan peratusan nilai skor mengikut metrik penilaian dalam pemprosesan bahasa tabii untuk penghurai ayat. Aliran proses yang terlibat dalam setiap fasa kajian ini ditunjukkan dalam Jadual 1.1.

Jadual 1.1

Reka Bentuk Kajian

REKA BENTUK KAJIAN		
Fasa	Aktiviti	Hasil
Mengategorikan dan analisis ayat untuk mendapatkan rumus binaan ayat	<ol style="list-style-type: none"> 1. Pengumpulan ayat 2. Kumpul atribut perkataan 3. Pengesahan lakaran 	<ol style="list-style-type: none"> 1. Ayat penyata BM bagi perkataan kurang atau sama 14 patah perkataan 2. Rumus yang disahkan 3. Atribut perkataan dan pangkalan data
Reka bentuk	<ol style="list-style-type: none"> 1. Model atribut perkataan 2. Model VPS dengan output tambahan 3. Algoritma penyemak dengan cadangan pembetulan struktur ayat 4. Algoritma VPS 5. Algoritma VPS dengan ouput tambahan 	<ol style="list-style-type: none"> 1. Model Atribut perkataan 2. Model VPS yang disahkan 3. Algoritma VPS dengan ouput tambahan
Pembangunan	<ol style="list-style-type: none"> 1. Reka bentuk antara muka 2. Hubungkan dengan pangkalan data 3. Pengaturcaraan setiap turutan algoritma 	Prototaip BMTutor
Penilaian dan rumusan	<ol style="list-style-type: none"> 1. Uji kaji mengikut kaedah Parseval 2. Penilaian pengguna menggunakan soal selidik <i>Usefulness, Satisfaction, and Ease of use (USE)</i> 	<ol style="list-style-type: none"> 1. Nilai <i>recall</i> dan <i>precision</i> 2. Nilai min bagi penilaian kepuasan subjektif dan kognitif pengguna

1.6 Sumbangan Kajian

Output tambahan pohon sintaksis yang dicadangkan berupa semakan dengan cadangan pembetulan ayat, atribut perkataan dan VPS ayat contoh. Output tersebut memerlukan rumus binaan ayat, model dan algoritma VPS dengan output tambahan sebagai panduan dalam mencapai objektif kajian. Gabungan output tambahan tersebut menghasilkan sumbangan berupa 1) model VPS dengan output tambahan, 2) algoritma VPS dengan semakan dan cadangan pembetulan ayat, 3) rumus binaan ayat, dan 4) pengecaman

atribut perkataan. Setiap sumbangan dijelaskan mengikut penerima manfaat seperti di bahagian seterusnya.

1.6.1 Model VPS dengan Output Tambahan

Model VPS yang direka bentuk boleh digeneralisasikan untuk kajian berkaitan pemprosesan ayat yang lain. Pembangun aplikasi pemprosesan bahasa dapat menggunakan model tersebut untuk mereka bentuk sistem berkaitan seperti penyemak ayat, penyemak jenis golongan kata dan mesin terjemahan serta mengembangkannya kepada bahasa lain.

1.6.2 Algoritma VPS dengan Semakan serta Cadangan Pembedulan Ayat

Turunan algoritma VPS yang berstruktur, sistematik dan boleh diguna pakai dalam menghasilkan kajian berkaitan akan menyumbang kepada pengkaji dan pembangun aplikasi. Algoritma ini boleh ditambah keunikan lain seperti penggunaan teks yang lebih panjang dan korpus ayat yang lebih luas. Ia dapat membantu para pengkaji dari segi masa dan kos. Sebagai contoh, pengkaji menjalankan kajian untuk menghasilkan teknik penyemak sintaksis Bahasa Inggeris (BI), mereka dapat menggunakan algoritma yang direka bentuk dengan mengubah struktur tatabahasa.

Semakan dengan cadangan pembedulan ayat merupakan sumbangan output baharu dalam bidang pengkomputeran linguistik. Bagi ayat yang didapati tidak sepadan dengan rumus yang disimpan, maka semakan bagi memaparkan cadangan pembedulan akan dilakukan. Sumbangan ini boleh dimanfaatkan oleh pengkaji dan pembangun aplikasi berasaskan pemprosesan bahasa tabii untuk dimajukan. Sebagai contoh, semakan dengan cadangan

pembetulan ke atas ejaan perkataan yang salah atau susunan penggunaan tatabahasa yang tidak tepat dalam aplikasi huraian teks.

1.6.3 Rumus Binaan Ayat

Rumus yang digunakan oleh pengkaji sedia ada dalam pemprosesan BM sememangnya memberi fokus kepada RSF. Namun setelah RSF diakui sebagai rumus tidak universal, maka rumus *X-bar* digunakan. Rumus ini telah digunakan dalam penghuraian ayat BI, tetapi belum pernah diuji dalam VPS ayat BM. Oleh yang demikian, kajian berkenaan rumus *X-bar* untuk kegunaan VPS bagi ayat penyata BM adalah sumbangan baharu yang diketengahkan dalam kajian ini.

Rumus BM *X-bar* yang dihasilkan dalam kajian ini untuk kegunaan VPS boleh membantu pengkaji bahasa dan pembina sistem untuk menghasilkan aplikasi lain yang berasaskan pemprosesan bahasa. Selain itu, rumus yang dicadangkan boleh diteruskan untuk menghasilkan rumus bagi ayat yang lebih kompleks. Rumus ini juga menyumbang kepada bidang linguistik untuk dimajukan dan dikembangkan.

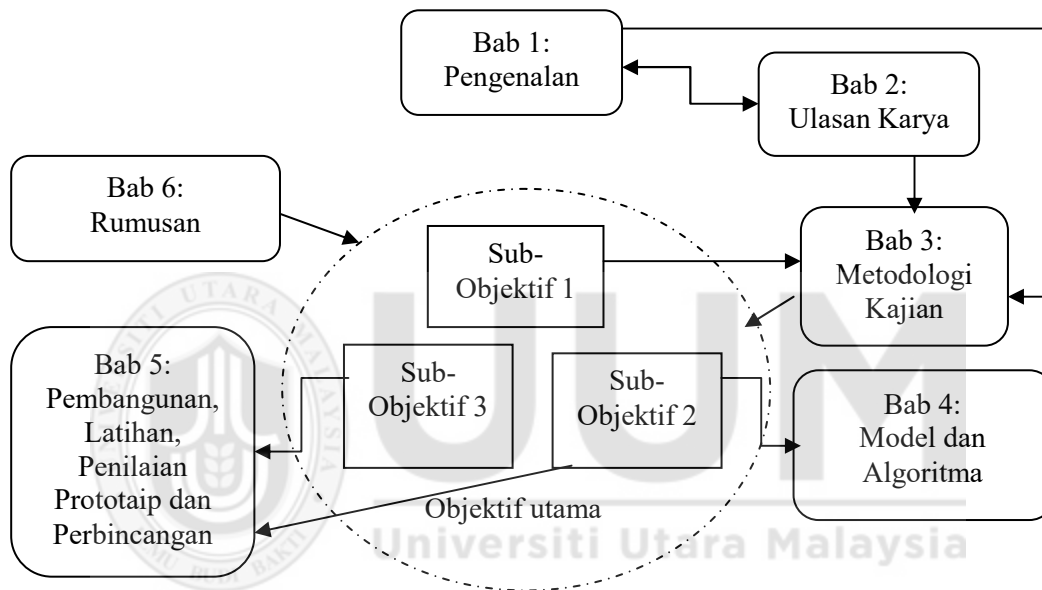
1.6.4 Pengecaman Atribut Perkataan

Satu ayat terdiri daripada frasa dan perkataan. Setiap perkataan mempunyai atributnya sendiri. Atribut ini berupa kelas kata, kata terbitan, terjemahan, imej dan ayat contoh. Penambahan atribut ini dalam VPS bertujuan membantu pemahaman ke atas setiap perkataan dengan lebih baik dari segi konsep ayat yang lain dan jenis perkataan sealiran. Atribut ini boleh dijadikan asas untuk melahirkan atribut lain seperti sebutan perkataan,

perkataan sinonim, dan contoh perkataan yang boleh digabung bagi membentuk frasa atau ayat yang lain.

1.7 Struktur Tesis

Tesis ini mengandungi enam bab. Ringkasan setiap bab dan perkaitan yang terlibat dalam mencapai objektif kajian ditunjukkan dalam Rajah 1.2.



Rajah 1.2. Struktur tesis berdasarkan objektif

Bab satu menerangkan tentang latar belakang kajian yang meliputi pernyataan masalah, persoalan, objektif, skop, sumbangan dan rangka kerja kajian. Penjelasan tersebut dijadikan panduan untuk Bab 2 hingga Bab 6.

Bab dua memberi fokus kepada lima perkara iaitu 1) latar belakang kajian, 2) teori yang mendasari kajian, 3) sorotan kritikal karya terdahulu berdasarkan persoalan kajian, 4)

pencapaian terkini karya terdahulu yang menyumbang kepada jurang kajian, dan 5) sumbangan yang hendak dilakukan. Penjelasan bab dimulakan dengan penerangan ringkas tentang latar belakang kajian diikuti penjelasan tentang teori iaitu teori *X-bar*, teori graf, teori gestalt and teori beban kognitif. Seterusnya, karya terdahulu dianalisis secara kritikal untuk mendapatkan komponen dan kaedah yang digunakan selain untuk menonjolkan keunikan kajian ini.

Bab tiga menjelaskan metodologi kajian dalam lima fasa. Fasa pertama adalah untuk mengenal pasti pernyataan masalah yang membawa kepada penentuan objektif dan skop kajian. Fasa kedua adalah untuk mengkategorikan dan menganalisis struktur ayat bagi mendapatkan rumus binaan ayat yang tepat. Analisis ini membawa kepada penggunaan rumus *X-bar*. Fasa ini juga menentukan kaedah binaan model, kaedah binaan algoritma, dan kaedah penentusahan model bagi kegunaan fasa ketiga iaitu fasa pembangunan prototaip yang berdasarkan kepada rumus, model dan algoritma. Seterusnya fasa penilaian menggunakan dua kaedah iaitu kaedah Parseval dan penilaian pengguna sebelum fasa terakhir memberi rumusan ke atas pencapaian objektif secara keseluruhan.

Bab empat memberi fokus kepada pembangunan model VPS dan algoritma. Penjelasan bab dimulakan dengan proses pembangunan model iaitu model atribut perkataan, dan model VPS dengan output tambahan. Model ini ditentusahkan sebelum diteruskan untuk menghasilkan algoritma pakej gabungan antara output tambahan. Menggunakan model VPS yang direka bentuk, algoritma VPS dengan output tambahan dihasilkan.

Bab lima bermatlamat untuk membuktikan konsep dalam model VPS yang dibina. Kaedah pembuktian dibuat melalui prototaip VPS. Latihan ke atas penggunaan prototaip dibuat sehingga dapat mengaplikasikan rumus *X-bar* dengan tepat bagi kegunaan fasa penilaian. Hasil penilaian prototaip dibincangkan dalam bentuk peratusan *recall* dan *precision*. Selain itu, penilaian pengguna turut dilakukan untuk mendapatkan peratusan skor min kepuasan subjektif dan kognitif pengguna.

Seterusnya, bab enam membuat rumusan keseluruhan berdasarkan sumbangan dan pencapaian objektif kajian. Bab ini diakhiri dengan memberi saranan ke atas kekangan yang dikenal pasti dan penambahbaikan yang boleh dilakukan supaya kajian ini dapat diteruskan atau ditambah baik untuk kegunaan kajian berkaitan di masa hadapan.

1.8 Rumusan Bab Satu

Objektif utama kajian adalah untuk mereka bentuk model VPS dengan output tambahan dan algoritma bagi pembangunan VPS tersebut. Asas kepada pembangunan model dan algoritma tersebut adalah rumus binaan ayat yang tepat dan lasak. Pencapaian objektif ini dapat dilakukan dengan membahagikan sub-objektif kepada tiga iaitu untuk, 1) mengkategorikan dan menganalisis ayat BM untuk mendapatkan rumus binaan ayat, 2) Mencadangkan model, algoritma dan prototaip kajian dan 3) Menjalankan uji kaji berdasarkan metrik penilaian dalam pemprosesan bahasa tabii untuk menguji ketepatan output dan membuat penilaian pengguna bagi mencapai kepuasan subjektif dan penerimaan kognitif pengguna.

Objektif kajian yang ditetapkan tertakluk kepada skop domain kajian, skop struktur ayat dan komponen dalam VPS. Secara keseluruhan, skop kajian ini adalah berlandaskan kepada domain BM yang mengambil bahan bacaan buku teks BM tingkatan satu hingga tingkatan lima bagi ayat kurang atau sama dengan 14 perkataan sebagai data kajian.

Sumbangan kajian menyumbang kepada pengkaji, pembangun aplikasi dan bidang pengkomputeran linguistik untuk dimajukan. Sumbangan ini berbentuk model VPS dengan atribut tambahan, algoritma VPS dengan semakan berserta cadangan pembetulan ayat, rumus binaan ayat dan pengecaman perkataan. Sumbangan tersebut mempunyai kepentingan tersendiri yang boleh dikembangkan secara berasingan atau secara gabungan menggunakan algoritma yang diperkenalkan dalam kajian ini.

Berpandukan kepada pemahaman konsep kajian yang merangkumi objektif, persoalan, skop, reka bentuk kajian dan ringkasan metodologi, kajian ini diteruskan untuk memahami dan mendalami secara kritis ulasan karya yang terlibat. Oleh yang demikian, bab ini dijadikan panduan untuk menjelaskan Bab Dua seterusnya.

BAB DUA

ULASAN KARYA

2.0 Pengenalan

Teknik visualisasi struktur maklumat berbentuk hierarki adalah kaedah untuk menggambarkan maklumat secara berstruktur yang dapat dibahagikan kepada tiga kategori iaitu secara senarai, garis dan diagram pohon (Wang, Wang, Dai, Wang, 2006). Antaranya seperti katalog produk, dokumen HTML, fail komputer dan carta organisasi. Kaedah ini dikenali sebagai rajah berpokok atau lebih dikenali sebagai pohon sintak atau dalam teori graf dikenali sebagai pokok berhierarki (Nguyen & Huang, 2002).

Banyak kajian atau alatan yang dihasilkan untuk menghasilkan visualisasi pokok berhierarki yang merangkumi pelbagai domain. Salah satu yang sangat dikenali adalah alatan Treemap yang dihasilkan oleh Universiti Maryland yang pertama kali direka bentuk oleh Ben Shneiderman pada tahun 1990 (University of Maryland, 2003). Antara contoh lain seperti Cone Trees, Hyperbolic Tree, 3D Hyperbolic Space, SpaceTree, dan Zoomology (Rusu, Santiago & Jianu, 2007), serta radial view dan disk tree (Nguyen & Huang, 2002). Pokok berhierarki ini digunakan dalam menghuraikan bahasa tabii dalam bidang pengkomputeran linguistik yang lebih dikenali sebagai pohon penghurai atau pohon sintaksis iaitu kaedah menghurai sintaksis bahasa.

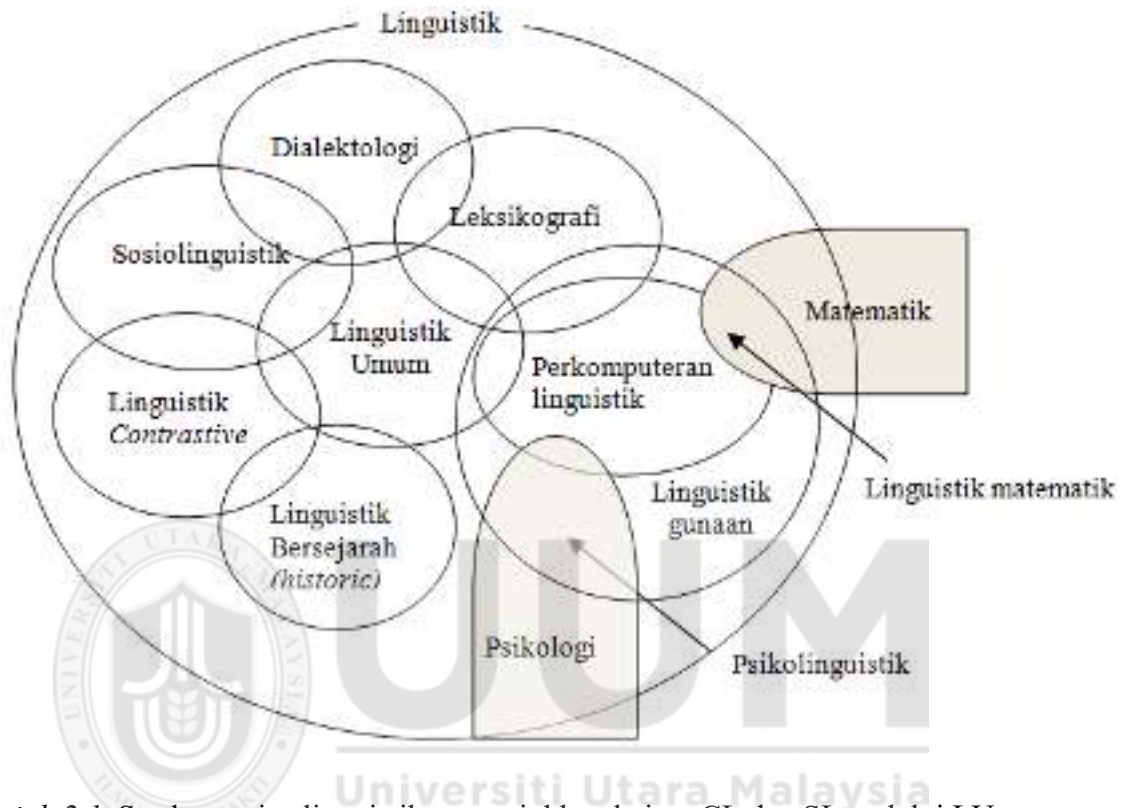
Latar belakang pengetahuan mengenai bidang ini serta ulasan karya terlibat perlu dianalisis agar objektif kajian dapat dicapai. Oleh itu, bab kedua ini memberi penjelasan mengenai ulasan karya dalam bidang pohon sintaksis yang bermatlamat untuk mendapatkan jurang yang boleh ditambah baik. Bab ini akan memberi fokus kepada lima perkara iaitu 1) latar belakang kajian, 2) teori yang mendasari kajian, 3) sorotan kritikal karya terdahulu berdasarkan persoalan kajian, 4) pencapaian terkini karya terdahulu yang menyumbang kepada jurang kajian, dan 5) sumbangan yang hendak dilakukan.

2.1 Latar Belakang Kajian

Latar belakang kajian menunjukkan kaitan antara bidang pengkomputeran linguistik dan sosiolinguistik melalui bidang linguistik umum. Bidang pengkomputeran linguistik ditakrif sebagai bidang interdisiplin untuk pemprosesan bahasa tabii. Bidang ini merupakan kombinasi antara sains komputer dan linguistik untuk mencapai taraf suatu sains. Bidang ini juga berkait rapat dengan bidang linguistik umum (Nederhof & Satta, 2013; Musthofa, 2010; Mohd Juzaidin, 2007; Mitkov, 2004; Bolshakov & Gelbulk, 2004; Zaharin, 1998). Manakala, bidang sosioinguistik pula adalah bidang kajian bahasa yang berkaitan dengan masyarakat (Abdul Razif & Rosfazila, 2016).

Perkaitan antara bidang pengkomputeran linguistik (CL) dan sosiolinguistik (SL) adalah melalui linguistik umum (LU) sebagaimana yang ditunjukkan dalam Rajah 2.1. Linguistik umum berkaitan dengan kajian mengenai fonologi, morfologi, sintaksis, semantik, dan pragmatik. Contoh kajian adalah berkenaan pemprosesan morfologi,

penghurai sintaksis, penghurai semantik, pengestrakan maklumat dan resolusi anafora (Mooney, 2004). Perkaitan ini ditunjukkan dalam Rajah 2.1 berikut.



Rajah 2.1. Struktur sains linguistik menunjukkan kaitan CL dan SL melalui LU.

(Sumber: Bolshakov & Gelbulk, 2004; Musthofa, 2010)

Rajah 2.1 menunjukkan bidang SL berkaitan dengan bidang CL yang mengkaji berkenaan LU. Oleh kerana sintaksis adalah kajian dalam bidang linguistik umum, maka rumus yang mendasari pembentukan sintaksis ini perlu dirujuk. Oleh itu, rumus pembentukan sintaksis ayat BM berbentuk rumus *X-bar* digunakan dalam menganalisis ayat BM.

Kajian berkenaan CL telah lama dikaji dalam pelbagai bahasa terutama di Malaysia. Sebagai contoh, kajian berkenaan penghurai ayat BM oleh Noor Hafhizah (2011) dan Ahmad Izuddin *et al.* (2007). Kajian tersebut menghasilkan pohon sintaksis sebagai output. Selain itu, banyak kajian lain yang dilakukan ke atas pemprosesan BM seperti kajian mengekstrak teks, morfologi dan penyemak ayat. Oleh itu, latar belakang kajian yang mendasari kajian pemprosesan ayat BM ini dijelaskan dalam bahagian seterusnya.

2.1.1 Kajian Pemprosesan Ayat di Malaysia

Di Malaysia, kajian pengkomputeran linguistik bermula pada tahun 1980-an mengenai penganalisis morfologi. Kajian ini dimulai oleh Zaharin Yusuf, Tengku Mohd Tengku Sembok dan Ahmad Zaki Abu Bakar (Mohd Juzaidin, 2007). Sejak dari itu pelbagai analisis tentang pengkomputeran tatabahasa dilakukan terutama di Universiti Sains Malaysia (USM) hingga tertubuhnya institusi terjemahan yang dikenali sebagai UTMK. Universiti Teknologi Malaysia dan Universiti Kebangsaan Malaysia juga menjalankan kajian dalam bidang ini (Zaharin, 2000).

Sehingga tahun 1990-an kajian dalam bidang pengkomputeran linguistik mula mendapat sambutan dengan terhasilnya penyemak ejaan, sistem perkamusan DBP, mesin terjemahan *Structured String Tree Correspondence* (SSTC) dan sebagainya. Selain itu, kajian peringkat kedoktoran juga dijalankan mengenai mesin terjemahan seperti Kong (1994) dan Zaharin (1986). Tahun 2000-an menyaksikan pembangunan aplikasi menjurus pelbagai sudut seperti di USM yang membangunkan pelayar internet BM dan kamus pelbagai bahasa (Chuah & Zaharin, 2002). Contoh lain seperti Norshuhani dan

Arina (2010) menghasilkan aplikasi ringkasan teks BM, penandaan kelas kata berdasarkan bahasa Arab oleh Jabar dan Tengku Mohd (2006) dan penandaan perkataan BM berdasarkan korpus Jawi (Juhaida, Khairuddin, Mohammad Faidzul & Mohd Zamri, 2016).

Kajian mengenai pemprosesan ayat juga mendapat galakan pengkaji di Malaysia. Antaranya seperti kajian untuk mengenalpasti persamaan ayat BM (Mohd Juzaidin, Fatimah, Abdul Azim, & Ramlan, 2008), pengsintesis ucapan ayat BM (Tan & Sh-Hussain, 2009), analisis sentimen automatik (Alsaffar & Nazlia, 2015), penyemak ayat BM (Rosmah, 1995; Suzaimah, 2002; Rozana *et al.*, 2011), penghurai ayat BM (Ahmad Izuddin *et al.*, 2007; Noor Hafhizah, 2011) dan alatan mengkategorikan teks BM (Maisarah, 2013). Antara alatan pemprosesan ayat, penghurai ayat atau sintaksis banyak diperlukan dalam menyokong alatan pemprosesan yang lain seperti mesin terjemahan, mengkategorikan ayat, penyemak, dan pengsintesis maklumat. Oleh itu, pemprosesan penghurai ayat ini dijelaskan dalam bahagian seterusnya.

2.1.2 Penghurai Sintaksis

Penghurai sintaksis melibatkan proses membuat pepadanan struktur sintaksis dalam ayat dengan tujuan untuk menghasilkan output berbentuk pohon sintaksis atau bentuk persembahan yang sesuai sebagai huraian ke atas ayat yang digunakan (Noor Hafhizah, 2011 dipetik daripada Jurafsky *et. al*, 2000). Penghuraian ini memerlukan rumus binaan ayat bagi bahasa yang dikaji (Tayal, Raghuwanshi & Malik, 2014).

Penghurai sintaksis dibahagikan kepada dua jenis iaitu penghurai statistik dan penghurai umum. Penghurai statistik digunakan oleh pengkaji yang bertujuan untuk mengurangkan keaburan struktur ayat. Antaranya seperti penghurai statistik BI (Nelson, Punch & Donaldson, 2011), penghurai wacana statistik (Soricut & Marcu, 2003), penghurai tanpa perkamusan (Klein & Manning, 2003), penghurai ayat bahasa Myanmar (Thant, Htwe & Thein, 2012), bahasa Rusia (Potemkin, 2009) dan bahasa Korea (Park & Kwon, 2008). Sumber rujukan utama kajian penghurai ayat statistik adalah berlandaskan kepada kajian penghurai statistik Charniak (2000) dan Collins (2000).

Penghurai sintaksis jenis umum adalah bertujuan untuk membuat semakan ayat berdasarkan rumus dan mengeluarkan output berbentuk pohon sintaksis. Antaranya penghurai ayat bahasa Arab (Shatnawi & Belkhouche, 2012; Shaalan, Farouk, & Rafea, 1999), penyemak sintaksis ayat BI (Tayal, Raghuwanshi, & Malik, 2014), dan kajian untuk mengekstrak teks subjektif yang menghasilkan output pohon sintaksis (Erfan & Lili, 2014). Penghasilan pohon sintaksis seperti kajian yang dinyatakan adalah berlandaskan kepada teori mengikut objektif yang hendak dicapai.

2.2 Kerangka Teori

Teori yang terlibat dalam kajian ini adalah teori *X-bar*, teori graf, teori gestalt dan teori beban kognitif. Teori graf digunakan untuk rujukan penghasilan VPS dan teori visualisasi maklumat seperti teori gestalt dan teori beban kognitif digunakan sebagai rujukan visualisasi atribut perkataan dan ayat. Rasional penggunaan teori tersebut dijelaskan dalam bahagian berikut.

Teori *X-bar*

Bab satu (skop domain) kajian ini telah menjelaskan bahawa, domain kajian ini adalah meliputi bahan bacaan pelajar sekolah menengah tingkatan satu hingga tingkatan lima untuk buku teks BM. Buku teks yang digunakan adalah buku keluaran Dewan Bahasa dan Pustaka (DBP) yang berdasarkan kepada buku Tatabahasa Dewan. Selain itu, buku Tatabahasa Dewan merupakan buku yang dicadangkan oleh Kementerian Pendidikan Malaysia sebagai buku sumber guru BM di sekolah. Buku ini mengetengahkan teori tatabahasa transformasi generatif (TTG) dalam pembentukan perkataan dan struktur ayat BM. Namun, setelah teori TTG ditambah baik oleh Chomsky (1970;1986), teori *X-bar* diperkenalkan. Rumus dikekalkan ditambah beberapa syarat agar boleh digunakan dengan lebih meluas dan boleh digunakan dalam bidang pengkomputeran.

Selain daripada itu, teori dalam kajian penghurai ayat melibatkan penggunaan tatabahasa formal seperti tatabahasa bebas konteks (CFG), tatabahasa kebergantungan (*dependency grammar*) atau tatabahasa lain yang bersesuaian dengan jenis output penghurai yang diskopkan (Nederhof & Satta, 2013). Teori *X-bar* adalah teori yang juga berdasarkan kepada CFG (Ramli, 1995).

Teori graf

Teori graf dirujuk dalam pembentukan pohon sintaksis kerana pohon sintaksis adalah salah satu contoh graf berhierarki.

Teori Gestalt dan Teori Beban Kognitif

Teori ini digunakan kerana pengkomputeran linguistik termasuk dalam bidang psikolinguistik. Teori yang terlibat dalam psikolinguistik adalah teori gestalt dan teori kognitif yang digunakan untuk menghuraikan komponen visualisasi kajian ini.

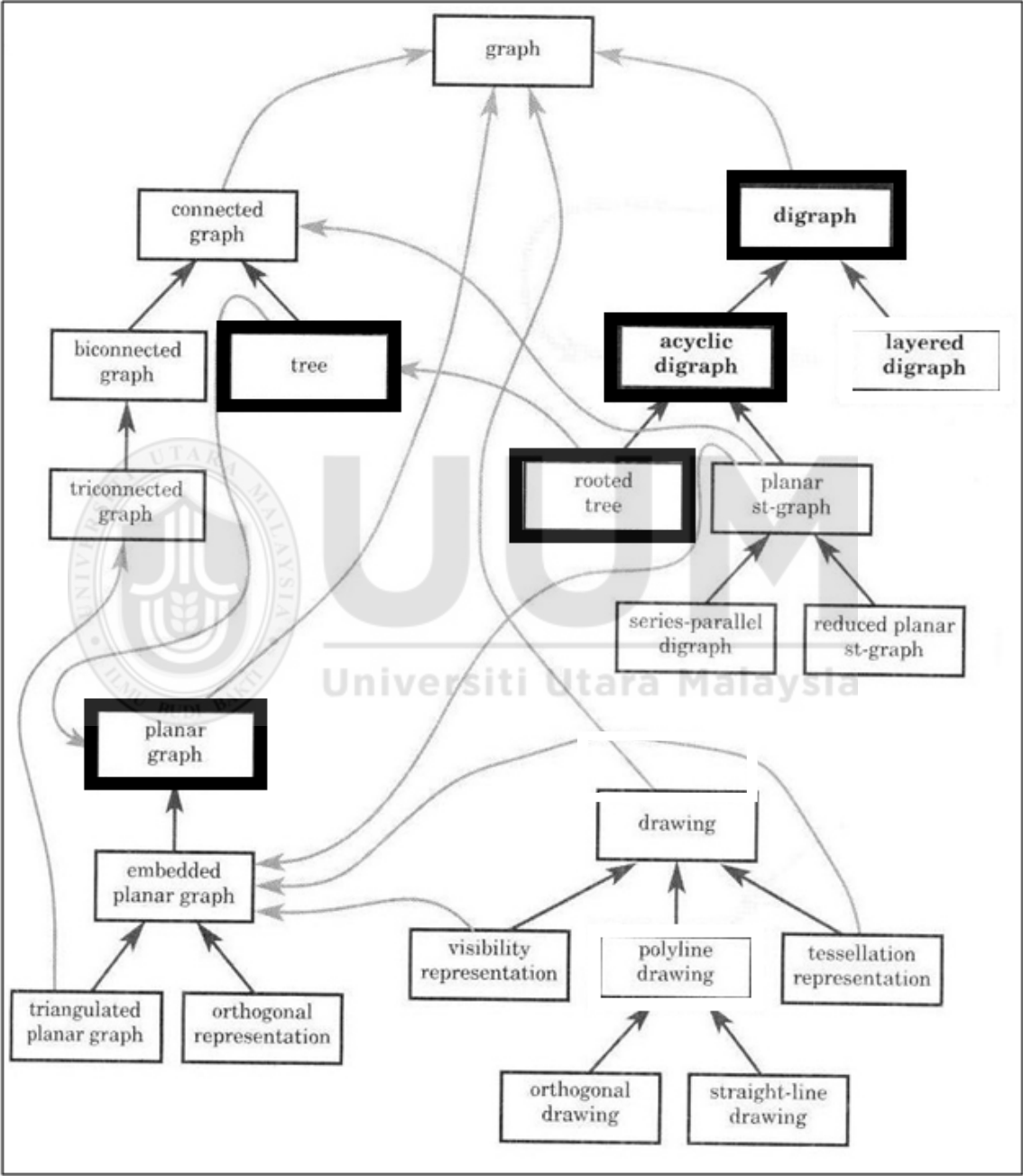
2.2.1 Teori Graf

Graf adalah struktur abstrak yang digunakan untuk memodelkan maklumat. Ia digunakan untuk mempersembahkan maklumat dalam bentuk objek bersambung. Oleh sebab itu, banyak sistem visualisasi maklumat memerlukan graf untuk melakar maklumat bagi memudahkan mereka membaca dan memahami (Battista, Eades, Tamassia & Tollis, 1999).

Graf mengandungi nod dan anak panah. Ia digunakan sebagai alatan visualisasi dalam pelbagai bidang untuk menyampaikan sesuatu maklumat supaya mudah difahami berbanding hanya melibatkan teks. Pohon sintaksis adalah salah satu jenis graf. Graf pula sebagai alatan visualisasi yang mempunyai nod dan anak panah. Untuk menghasilkan VPS yang baik seperti yang diskopkan, maka teori graf perlu diberi perhatian. Hal ini bagi memahami struktur pembentukan graf pohon sintaksis yang menepati skop kajian dan dapat menghuraikan ayat BM seperti yang diperlukan.

Menurut Battista *et al.* (1999), dalam melakar gambaran graf terdapat beberapa pendekatan berbeza yang digunakan dalam bidang yang berbeza. Antaranya pendekatan berhierarki, *visibility*, tambahan, *force-directed*, dan *divide dan conquer*. Dalam kajian ini, pendekatan berhierarki dipilih berdasarkan kepada Skop VPS dalam Bab Satu. Graf

juga dapat dibahagikan kepada beberapa jenis iaitu *digraph*, *connected graph* dan *planar graph* seperti dalam Rajah 2.2. Petak yang dihitamkan menunjukkan aliran graf berbentuk hierarki yang difokuskan dalam kajian ini.



Rajah 2.2. Pendekatan graf berhierarki

(Sumber: Battista *et al.*, 1999)

Rajah 2.2 menunjukkan pohon sintaksis dalam kajian ini dipanggil sebagai *rooted tree* (pohon berakar) kerana nod dihasilkan daripada *root* atau akar yang bermula daripada atas. *Rooted tree* adalah salah satu graf *acyclic digraph* yang mempunyai lakaran berbentuk *planar* iaitu gambaran anak panah yang tidak mempunyai penyimpangan antara anak panah yang lain.

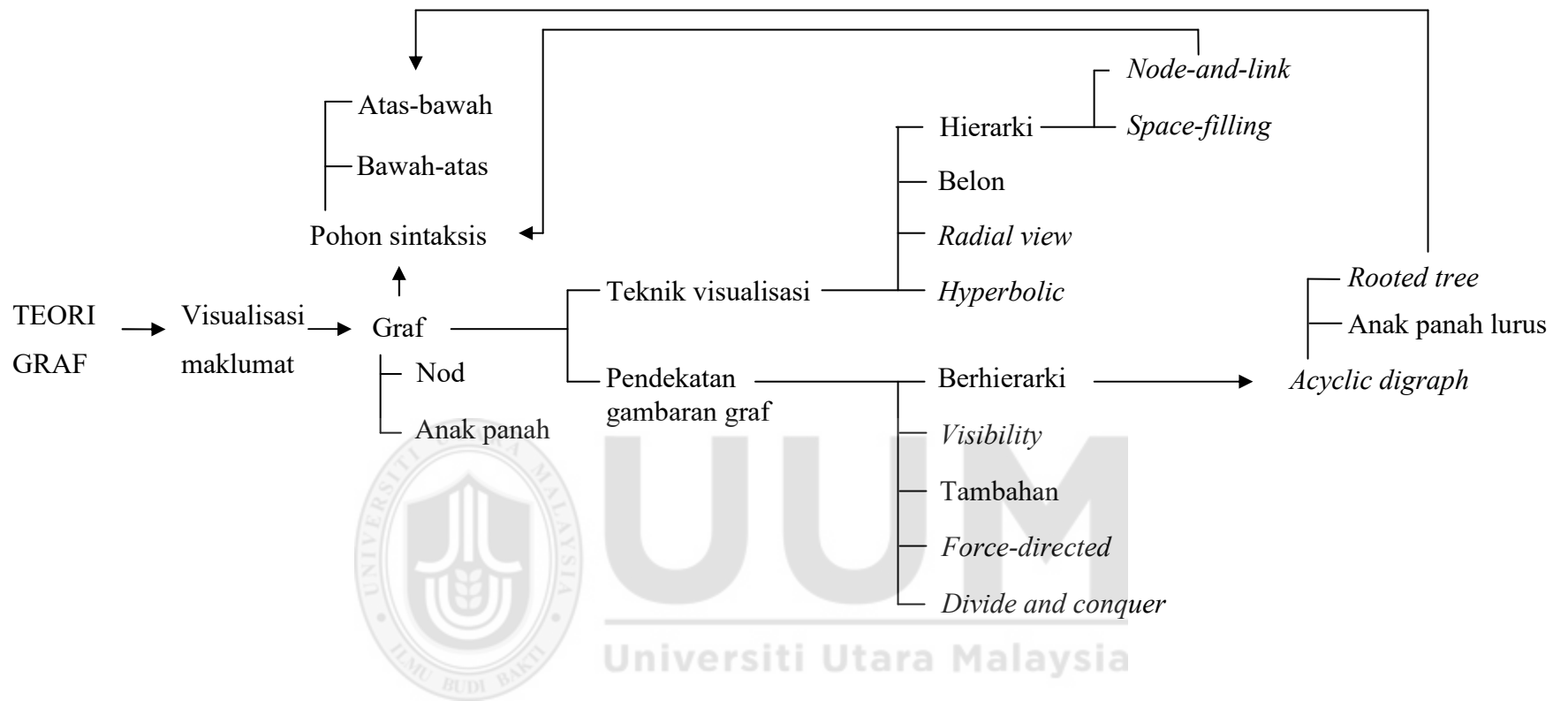
Oleh itu, dapat disimpulkan bahawa, pohon sintaksis yang difokuskan adalah dalam kategori pohon sintaksis berhierarki dan bersambung (*acyclic digraph*) antara anak panah dan nod. Pohon sintaksis juga mempunyai akar (*rooted tree*) yang menyambungkan nod atas dengan nod bawahan menggunakan anak panah berbentuk lurus dan tidak menyimpang (*planar*) antara anak panah yang lain, seperti keperluan dalam Skop VPS yang dijelaskan dalam Bab Satu.

Penghuraian maklumat berhierarki melibatkan dua kaedah berbeza. Pertama, kaedah *node-and-link diagram*, di mana sudut graf tersebut dipersembahkan dengan menggunakan garisan. Kedua, kaedah *space-filling* yang memaparkan struktur maklumat dengan cara persembahan nod visual secara bersarang atau dengan kebergantungan persekitaran (Johnson & Shneiderman, 1991; Luboschik, & Schumann, 2007). Kaedah *space filling* adalah kaedah yang banyak diberi perhatian seperti *treemaps*, *Grokker*, dan *nested circles*. Ia memberi fokus kepada pendekatan berdasarkan ruang untuk visualisasi struktur maklumat secara hierarki dan tidak melibatkan penggunaan nod dan anak panah.

Daripada dua kaedah tersebut, terdapat pelbagai teknik visualisasi yang boleh digunakan seperti teknik hierarki, teknik belon (*balloon view*), teknik *radial view*, dan teknik *hyperbolic*. Dalam menggambarkan struktur bahasa, teknik hierarki adalah teknik yang paling sesuai digunakan kerana gambaran perkataan dibuat secara jujukan atas-bawah. Teknik berhierarki juga menyokong tugas yang berasaskan *label* atau *atribut* (Lee, 2006), yang digunakan dalam pohon sintaksis seperti dalam kajian ini. Selain itu, struktur pohon sintaksis yang dihasilkan berbentuk nod dan sub-nod yang juga sesuai untuk memaparkan struktur tatabahasa bagi ayat yang hendak dipaparkan.

Secara keseluruhan, Rajah 2.3 menunjukkan perkaitan antara komponen teori graf yang berkaitan dengan skop kajian ini.





Rajah 2.3. Perkaitan teori graf dengan skop kajian

Merujuk kepada Rajah 2.3, seperti yang telah dijelaskan, VPS adalah kajian berkaitan dengan visualisasi maklumat. Kaedah penerangan visualisasi maklumat boleh dilakukan dengan pelbagai cara dan penerangan menggunakan graf memang seringkali digunakan. Penerangan ini melibatkan nod dan anak panah. Kaedah nod dan anak panah dalam teknik hierarki adalah kaedah persembahan pohon sintaksis. Selain daripada itu, untuk memaparkan pohon sintaksis daripada binaan ayat, teori *X-bar* dijelaskan dalam bahagian seterusnya.

2.2.2 Teori *X-bar*

Pada asasnya teori *X-bar* menekankan prinsip bahawa setiap frasa perlu mengandungi kepala (*head*) yang unik. Teori ini ditambah baik dari teori TTG (Ramli, 1995) dengan syarat bahawa setiap nod hanya bercabang dua (Mazura, 2002; Nasrun, 1994) dan frasa mesti mempunyai kepala iaitu X. Kepala X akan mempunyai maksimal X-frasa iaitu XP dan mempunyai frasa pertengahan yang dipanggil sebagai X' (disebut sebagai *X-bar*) (Jubilado, 2010).

Sintaksis dalam teori *X-bar* berkaitan dengan teori graf apabila sintaksis menjadi kaedah atau data yang dipersembahkan dalam keperluan pohon sintaksis. Pohon sintaksis digunakan untuk membuat huraian tentang maklumat dalam ilmu bahasa. Pohon sintaksis lahir daripada teori graf dan ilmu bahasa pula berkaitan dengan teori *X-bar*. Dari segi visualisasi pula, terdapat banyak teori yang boleh digunakan. Antara teori yang berkait rapat dengan visualisasi mengikut kaedah psikologi adalah teori gestalt dan teori beban kognitif (Erfan & Lili, 2014). Oleh yang demikian, keperluan teori ini dibincangkan dalam bahagian seterusnya.

2.2.3 Teori Gestalt

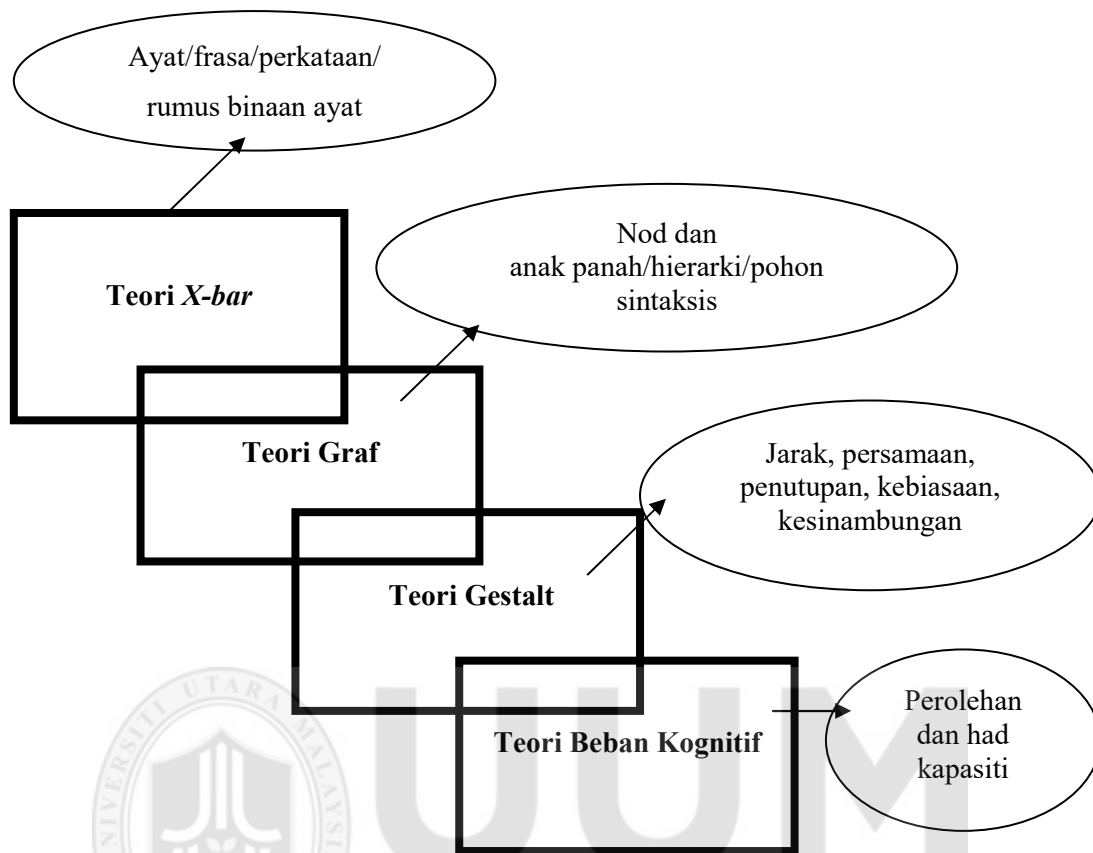
Visualisasi telah digunakan dalam memahami ilmu linguistik (Zhao, Chevalier, Collins, & Balakrishnan, 2012). Visualisasi membantu kebolehan manusia untuk memahami (Grinstein & Ward, 2002) melalui aktiviti kesedaran manusia. Ia adalah pengalaman visual dalam melihat data yang dipersembahkan dalam paparan antara muka (Spence, 2007). Oleh itu, teori yang perlu dipertimbangkan dalam reka bentuk antara muka dalam paparan visualisasi adalah teori gestalt (Hicks, 2009).

Teori ini mempertimbangkan kedudukan sesuatu benda yang terdapat dalam paparan visualisasi (Erfan & Lili, 2014). Antara sembilan prinsip yang boleh diikuti adalah *pragnanz* (pengamatan), *proximity* (jarak), *similarity* (persamaan), *symmetry* (simetri), *closure* (penutupan), *continuity* (kesinambungan), *common fate*, *familiarity* (kebiasaan) dan *figure and ground*. Dalam paparan yang melibatkan nod dan anak panah, prinsip "*continuity*" boleh digunakan (Hicks, 2009). Ia bermaksud penggunaan susunan bersambung antara nod secara lurus. Empat prinsip jarak, persamaan, penutupan, dan kebiasaan adalah prinsip yang sering ditekankan (Azizi, Asmah, Zurihanmi & Fawziah, 2005). Kaedah penyusunan bahan dalam antara muka adalah untuk membantu penerimaan pengguna yang dikenali sebagai persepsi. Ini bermaksud, jika bahan dapat disusun dengan baik, maka penerimaan juga akan mudah dilakukan (Ware, 2013). Gaya penerimaan dan persepsi ini adalah berkaitan dengan kognitif pengguna.

2.2.4 Teori Beban Kognitif

Potensi pembelajaran seseorang individu dipengaruhi oleh gaya kognitif dan cara maklumat dipersembahkan (Ahmad Rizal & Yahya, 2006). Apabila mengaplikasikan pendekatan kognitif dalam pembelajaran, tahap kefahaman perlu difokuskan (Azizi *et al.*, 2005). Kefahaman terhadap bahan pembelajaran boleh diterap dengan menghasilkan modul atau aplikasi berasaskan teori beban kognitif (Sun, Zaidatun & Jamalludin, 2007).

Teori beban kognitif melibatkan komponen skema perolehan dan had kapasiti. Teori ini boleh digunakan dalam reka bentuk pembelajaran dengan objektif untuk mengambil kira kebolehan dan kekangan pemprosesan maklumat. Pemahaman dan penerimaan berkait rapat dengan teori ini dengan mengambil kira kaedah paparan pembelajaran yang dipersembahkan (Plass, Moreno, & Brunken, 2010). Oleh itu perolehan yang dimaksudkan berkaitan dengan penerimaan dan kefahaman dengan had kapasiti penerimaan tidak membebankan pengguna. Teori ini berkaitan dengan teori gestalt kerana teori gestalt membantu persembahan paparan supaya mudah diterima dan difahami. Perkaitan antara teori ini ditunjukkan dalam Rajah 2.4.



Rajah 2.4. Kerangka teori

Rajah 2.4 menunjukkan teori *X-bar* dan teori graf berkaitan kerana kedua-dua teori membincangkan tentang pohon sintaksis. Pohon sintaksis dibentuk daripada ayat binaan daripada rumus dalam teori *X-bar* dan gambaran graf pohon sintaksis menggunakan teori graf. Teori gestalt pula berkaitan dengan teori graf apabila tahap kesinambungan mempunyai persamaan dengan komponen anak panah dalam teori graf. Selain itu, prinsip teori gestalt juga membantu pemahaman kognitif pengguna dengan menggunakan prinsip kesinambungan dan persamaan dalam reka bentuk dan susunan nod, anak panah dan paparan VPS. Seterusnya, setelah asas teori difahami, kajian diteruskan dengan menganalisis ulasan karya terlibat.

2.3 Kajian Berkaitan Rumus Binaan Ayat BM

Kajian pohon sintaksis sememangnya berkait rapat dengan penggunaan rumus binaan ayat terutama bagi ayat BM. Sebagai contoh, kajian Noor Hafhizah (2011), Ahmad Izuddin *et al.* (2007), Suzaimah (2002) dan Rosmah (1995) menggunakan rumus binaan ayat BM. Antara kajian tersebut, kajian Noor Hafhizah (2011) telah menerbitkan sebanyak 147 rumus bagi 1000 ayat mudah. Namun, kajian lain yang sealiran tidak menerbitkan rumus sebagaimana yang dihasilkan Noor Hafhizah (2011) kerana penyelidik terlibat menjadikan rumus asas Nik Safiah (1995;2009) sebagai rujukan. Namun, persoalannya adakah rumus binaan ayat seperti RSF boleh digunakan dalam VPS sebagaimana yang dilakukan oleh penyelidik terdahulu? Adakah rumus ini masih digunakan dalam VPS kini? dan bolehkah rumus ini ditambah baik untuk mendapatkan kaedah mengkategorikan struktur ayat dan mendapatkan rumus binaan ayat bagi kegunaan VPS? Persoalan ini perlu dijawab bagi mendapatkan jawapan kepada persoalan kajian yang pertama.

Penerbitan sebanyak 147 rumus dalam kajian Noor Hafhizah (2011) membuktikan bahawa rumus binaan ayat BM sememangnya boleh digunakan dalam VPS. Hal ini kerana kajian tersebut telah berjaya mengeluarkan output pohon sintaksis. Namun, para pengkaji bidang ilmu bahasa mengakui bahawa rumus binaan ayat seperti RSF mempunyai kelemahan yang boleh ditambah baik.

Perbincangan mengenai kelemahan RSF yang dianggap tidak universal dan bilangan tahap kategori yang terhad banyak dibincangkan oleh penyelidik bahasa setelah

Chomsky (1970;1986) mencadangkan penambahbaikan ke atas kekangan yang beliau sedari. Pada ketika ini teori *X-bar* dicadangkan, namun konsep rumus, transformasi dan kesejagatan masih dikekalkan. Selain itu, Mazura (2002), Jubilado (2010), Mohd Juzaidin (2008), Ramli (1995) dan Nasrun (1994) antara penyelidik yang turut memberi ulasan mengenai kekangan RSF yang membawa kepada kajian mengenai *X-bar* untuk ayat BM dan BI. Oleh yang demikian, rumus *X-bar* adalah rumus binaan ayat terkini yang boleh dikaji dan digunakan dalam VPS ayat BM. Namun, bolehkah rumus ini dizahirkan melalui pengaturcaraan dan boleh memaparkan VPS?

Noor Hafhizah (2011) dalam membangunkan prototaip penghurai ayat statistik BM telah mengumpul 1000 ayat mudah untuk perkataan kurang daripada 10 dari pelbagai sumber. Ayat tersebut dihantar kepada Munsyi Dewan untuk ditandakan kelas kata dan frasa. Terdapat sesetengah ayat diubah kedudukan perkataan supaya ayat mudah ditandakan. Kaedah yang digunakan dalam pengumpulan ayat dan rumus tidak dinyatakan dengan jelas, namun melalui hasil penulisan tesis sarjana yang dihasilkan dapat dinyatakan kaedah yang diikuti adalah pengumpulan ayat mudah, penandaan kelas kata oleh Munsyi Dewan dan pengumpulan rumus. Kaedah yang digunakan iaitu pengumpulan ayat, penandaan kelas kata dan kaedah pengumpulan rumus dapat dijadikan panduan.

Selain itu, komunikasi peribadi dengan Prof. Emeritus Datuk Dr. Nik Safiah Karim sebagai pakar BM telah dilakukan untuk mengetahui tentang rumus binaan ayat BM yang telah dihasilkan. Antara maklumat yang diperoleh adalah seperti berikut.

"Rumus dalam Tatabahasa Dewan itu tidak lengkap kerana hanya untuk memperkenalkannya sahaja. Saya ada membuat sedikit rumus dalam tesis Ph.D saya, yang telah diterbitkan oleh Dewan Bahasa dan Pustaka. Sila rujuk buku tersebut, kalau-kalau dapat diambil manfaat. Tetapi terlebih dahulu saya mengaku analisis itu agak ketinggalan zaman kerana tesis saya selesai pada akhir tahun 1976" (Nik Safiah Karim, komunikasi peribadi, Julai 12, 2011).

Oleh yang demikian, penghasilan rumus $X\text{-bar}$ bagi ayat penyata BM untuk kegunaan VPS menjadi salah satu sumbangan kajian ini. Kaedah mendapatkan rumus $X\text{-bar}$ melalui lakaran pohon sintaksis seperti yang dijelaskan dalam Ramli (1995) dan Mazura (2002) dijadikan panduan dalam lakaran pohon sintaksis. Manakala kaedah Noor Hafhizah (2011) iaitu pengumpulan ayat, lakaran pohon sintaksis dan pengumpulan rumus dijadikan panduan untuk mengkategorikan struktur ayat dan mendapat rumus yang diperlukan.

2.4 Kajian Berkaitan Model, Algoritma dan Prototaip

Bagi menjawab persoalan kedua kajian iaitu "Bagaimana model, algoritma dan prototaip output tambahan pohon sintaksis boleh dibina?", kajian ini diteruskan dengan terlebih dahulu memahami makna model dan algoritma. Jurafsky dan Martim (2009) menyatakan bahawa model mengandungi komponen, perkaitan antara komponen dan persembahan. Algoritma pula adalah sebarang prosedur perkomputeran yang jelas dan memerlukan beberapa nilai, atau set nilai, sebagai input dan menghasilkan beberapa nilai, atau set nilai, sebagai output dalam menyelesaikan masalah. Algoritma

menerangkan prosedur perkomputeran tertentu untuk mencapai perkaitan antara input dan output (Cormen, Leiserson, Rivest & Stein, 2001).

Matlamat bahagian ulasan ini adalah untuk mendapatkan komponen dan kaedah daripada karya terdahulu. Komponen tersebut boleh diperolehi melalui analisis kritikal ke atas model terlibat. Manakala kaedah pula memerlukan analisis terhadap kajian penghurai ayat bagi mendapatkan algoritma reka bentuk pohon sintaksis. Oleh itu, proses menganalisis ulasan karya ini adalah seperti berikut.

1. Model yang digunakan dalam sorotan karya yang terlibat adalah model SSTC.

Untuk menjelaskan konsep dalam pembangunan penghurai ayat, model dalam penghurai ayat secara sintaksis dan semantik juga dikaji. Terdapat dua kajian yang diambil sebagai rujukan iaitu model penghurai semantik dan model penghurai sintaksis.

2. Algoritma penghurai ayat dalam kajian terdahulu dikaji untuk mendapatkan kaedah penghuraian ayat. Perbandingan pohon sintaksis dilakukan mengikut skop kajian iaitu pohon sintaksis yang berbentuk konkrit atau *parse tree*, mempunyai nod dan anak panah, dan berhierarki.

Bagi mendapatkan komponen dan kaedah yang diperlukan, analisis kritikal dilakukan ke atas 19 kajian terlibat termasuk kajian dalam Lampiran A. Ulasan ke atas kajian tersebut dibahagikan kepada model, penghurai ayat, penyemak, cadangan, visualisasi dan atribut perkataan. Setiap karya ini disenaraikan dalam Jadual 2.1.

Jadual 2.1

Kajian Berkaitan

Kategori	Aplikasi/Kajian	Tahun	Bahasa
Model	1. <i>SSTC</i>	1998	BI
	2. Penghurai ayat	2006	BI
	3. Penghurai semantik	2008	Matematik
Penghurai ayat	4. <i>Lehner's Prolog Tree Drawing</i>	1994	BI
	5. <i>phpSintakTree</i>	2003	BI
	6. Penghurai Ayat Bahasa Melayu	2007	BM
	7. <i>SynView</i>	2009	BI
	8. <i>RSyntaxTree</i>	2012	BI, Jepun, China Korea
	9. Penghurai Ayat Bahasa Arab	2012	Arab
Penghurai statistik ayat	10. Penghurai statistik ayat Korea	2008	Korea
	11. Penghurai Statistik Ayat Bahasa Melayu	2011	BM
	12. Penghurai statistik ayat bahasa Myammar	2012	Myammar
Penghurai tatabahasa	13. <i>Link Grammar</i>	1991	BI
Penyemak ayat	14. Penyemak sintak	1995	BM
	15. Penyemak ayat BM	2002	BM
	16. Penyemak sintak ayat BI	2014	BI
	17. Ekstrak Visualisasi Teks Subjektif	2014	BI
Semakan dengan cadangan	18. Penyemak ayat BM	2011	BM
Atribut perkataan	19. <i>MALEX</i>	2010	BM

Kajian berkenaan penghurai ayat atau pohon sintaksis telah diperkenalkan sejak tahun 1980-an bagi banyak bahasa termasuk BM. Penghurai, penyemak, cadangan ayat dan

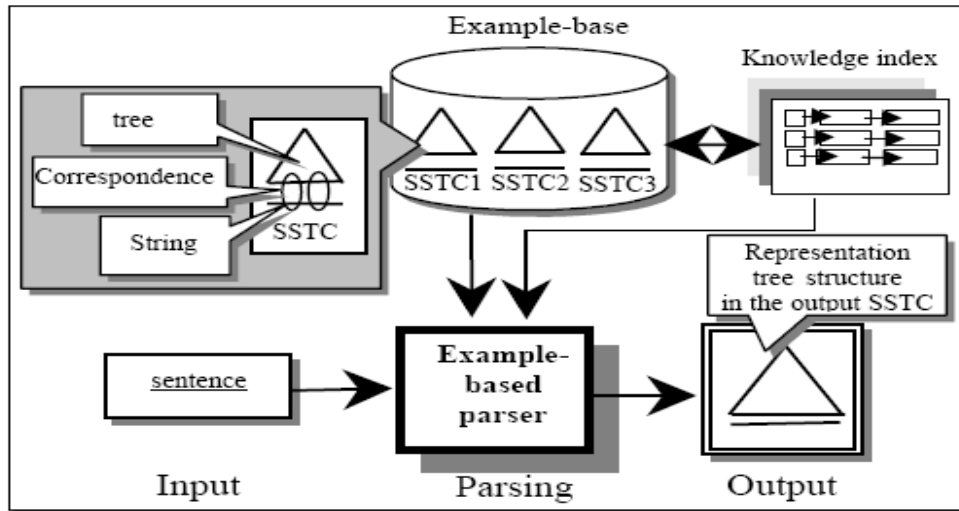
atribut perkataan sememangnya telah diperkenalkan dalam BM sejak tahun 2000-an. Sehingga tahun 2014, kajian bagi BM masih terus mendapat sambutan penyelidik dalam menghasilkan ekstrak visualisasi teks. Setiap kajian terlibat ini dianalisis dalam bahagian seterusnya.

2.4.1 Model

Model SSTC, model penghurai ayat dan model penghurai semantik adalah kajian yang berkaitan untuk dianalisis bagi mendapatkan komponen yang diperlukan dalam model VPS dengan output tambahan. Penjelasan ini dihuraikan dalam bahagian 2.4.1.1 hingga 2.4.1.3.

2.4.1.1 Model SSTC

Penganalisis sintak SSTC (Al-Adhaileh & Kong, 1998) telah dihasilkan bagi kegunaan mesin terjemahan BI ke BM berdasarkan contoh. Setiap contoh yang terdapat dalam pangkalan data yang dikenali sebagai *example base*, akan diberi simbol SSTC. Simbol tersebut merangkumi huraian tentang ayat, persembahan pohon, perkaitan antara sub-string ayat dan perkaitan antara sub-pohon. Dalam proses menganalisis struktur ayat, paparan pohon bagi setiap frasa yang terlibat adalah dengan menggunakan pendekatan bawah-ke-atas (*bottom-up*). Sub-pohon yang dihasilkan akan digabung untuk membentuk pohon yang juga akan berdasarkan kepada contoh menggunakan pendekatan atas-bawah (*top-down*). Model yang digunakan ditunjukkan dalam Rajah 2.5.



Rajah 2.5. Model penganalisis sintak SSTC

(Sumber: Al-Adhaileh & Kong, 1998)

Rajah 2.5 menunjukkan padanan dibuat dengan memecahkan ayat kepada frasa. Frasa yang hampir sama dengan frasa yang terdapat dalam pangkalan data dipadankan dengan struktur pohon di mana setiap frasa akan mempunyai struktur binaan pohon yang berbeza. Kaedah yang digunakan untuk membuat padanan adalah berdasarkan kepada indeks pengetahuan yang telah ditetapkan dalam pangkalan data.

Rajah 2.5 juga menunjukkan komponen yang terlibat dalam model pembangunan SSTC adalah **ayat**, **frasa**, **pangkalan data** (*example-base*) dan **penghurai ayat**. Gabungan penggunaan komponen tersebut berjaya menghasilkan sub-pohon bagi kegunaan mesin terjemahan. Walaupun penghasilan pohon yang dilakukan tidak mengikut padanan rumus ekor tujuan penghasilannya untuk kegunaan proses penterjemahan ayat, namun kaedah penghuraian dan komponen yang terlibat boleh membantu kajian ini dalam penentuan komponen yang diperlukan.

2.4.1.2 Model Penghurai Ayat

Murugesan dan Cassimatis (2006) menghuraikan tentang komponen yang diperlukan dalam sesebuah penghurai ayat. Menurut mereka, sesebuah penghurai ayat boleh membantu pemahaman dalam bidang yang dikaji iaitu mengenai penyatuan bahasa dalam kognitif. Komponen yang diperlukan dalam sesebuah penghurai ayat boleh mengandungi **perwakilan perkataan, perwakilan frasa, struktur ciri-ciri, rumus tatabahasa** dan **rumus leksikal**.

Walaupun kajian tersebut tidak menghasilkan aplikasi VPS mahupun sebuah lakaran model yang jelas, namun komponen yang disenaraikan iaitu perkataan, frasa, leksikal dan tatabahasa atau rumus adalah komponen asas sesebuah penghuraian ayat. Namun, komponen tersebut boleh dijadikan asas dalam menentukan komponen kajian ini.

2.4.1.3 Model Penghurai Semantik

Peters (2008) memperkenalkan model semantik yang digunakan dalam pembelajaran matematik. Penghurai ayat secara semantik digunakan untuk mendapatkan pemahaman dalam pembelajaran matematik. Menurut beliau, pembelajaran matematik melibatkan penggunaan bahasa tabii bagi memahami tentang tatabahasa. Pemahaman ini dapat diperoleh dengan cara memahami pakej penggunaan perkataan untuk mendapatkan maksud daripada turutan perkataan yang terlibat.

Model semantik yang dihasilkan adalah perkaitan antara komponen **penghurai ayat, leksikon** dan **tatabahasa**. Komponen ini adalah komponen asas dalam penghuraian ayat

seperti komponen dalam kajian SSTC dan penghurai sintaksis. Namun, komponen tersebut boleh dijadikan komponen asas dalam kajian ini.

2.4.1.4 Implikasi daripada Model Terdahulu

Ringkasan pembangunan model berpandukan sorotan karya yang dianalisis adalah seperti dalam Jadual 2.2.

Jadual 2.2

Ringkasan Sorotan Karya Tentang Pembangunan Model

Ciri-ciri	SSTC (Al-Adhaileh & Kong, 1998)	Model penghurai sintaksis (Murugesan & Cassimatis, 2006)	Model penghurai semantik (Peters, 2008)
Tujuan kajian	Untuk kegunaan mesin terjemahan BI ke BM berdasarkan contoh	Untuk kajian penyatuan bahasa dalam kognitif	Untuk mendapatkan pemahaman dalam pembelajaran matematik
Komponen model	1. Ayat 2. Frasa 3. Pangkalan data (<i>example-base</i>) 4. Penghurai ayat	1. Perwakilan perkataan 2. Perwakilan frasa 3. Struktur ciri-ciri 4. Rumus tatabahasa 5. Rumus leksikal	1. Penghurai ayat 2. Leksikon 3. Tatabahasa

Komponen yang diperlukan dalam pembangunan penghurai ayat mengandungi leksikon atau pangkalan data dan perwakilan perkataan dan frasa yang mewakili sesebuah ayat. Jadual 2.2 memberi rumusan ke atas komponen terlibat yang dapat diringkaskan seperti **perkataan/frasa/ayat, leksikon/ pangkalan data, penghurai ayat dan rumus tatabahasa/leksikal**. Walau bagaimanapun, komponen ini tidak mempunyai semakan,

cadangan, dan atribut perkataan. Komponen ini diambil sebagai rujukan seperti 1) perkataan/ayat, 2) penghurai ayat/ VPS, 3) rumus tatabahasa atau binaan ayat bagi mereka bentuk algoritma VPS seperti yang dijelaskan dalam bahagian seterusnya.

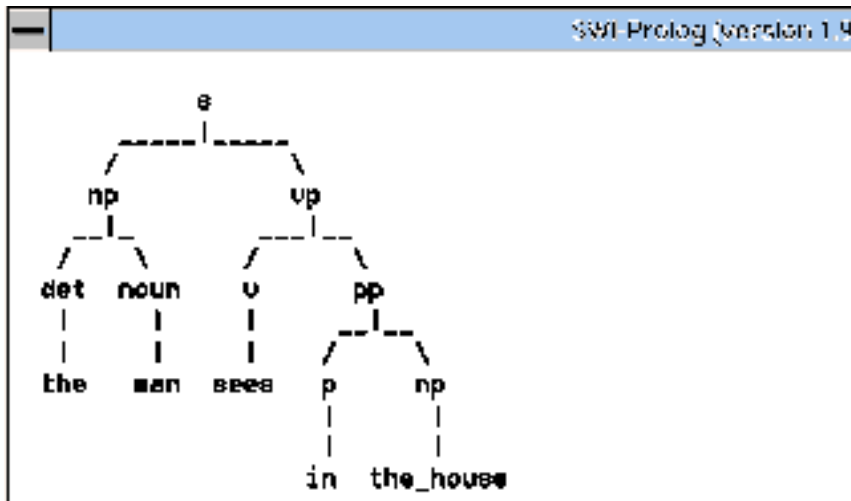
2.4.2 Algoritma Penghurai Ayat

Terdapat 10 kajian berkaitan dengan penghuraian ayat. Kajian ini adalah kajian penghurai Lehner, phpSintakTree, SynView, RSyntaxTree, penghurai ayat BM, penghurai ayat Bahasa Arab, penghurai statistik (BM, Myammar, Korea) dan penghurai tatabahasa Link Grammar.

Kajian tersebut dianalisis untuk mendapatkan kaedah penghuraian ayat yang dilakukan untuk pembangunan algoritma VPS dengan output tambahan. Setiap kajian dianalisis dan dijelaskan dalam bahagian seterusnya.

2.4.2.1 Penghurai Lehner's Prolog Tree Drawing

Penghurai yang direka bentuk untuk ayat BI pada tahun 1994 ini adalah penghurai untuk bahasa pengaturcaraan Prolog. Input yang diterima mesti mengikut struktur perisian Prolog iaitu memerlukan tanda koma bagi mengasingkan setiap perkataan dalam ayat. Walaupun kegunaannya boleh memberi kesukaran kepada pengguna yang tidak memahami perisian Prolog, namun kaedah paparan output dan cara pembahagian subjek-predikat boleh menyumbang idea untuk paparan VPS kajian ini seperti Rajah 2.6.



Rajah 2.6. Penghurai **Lehner's Prolog Tree Drawing**

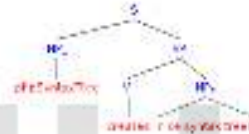
(Sumber: Dougherty, n.d)

Walaupun penghurai Lehner's Prolog Tree Drawing tidak melibatkan tambahan output lain selain paparan pohon sintaksis, namun kaedah paparan output tersebut boleh menyumbang idea dalam pembangunan prototaip kajian ini. Output yang dihasilkan menunjukkan sebuah ayat boleh dipecahkan kepada subjek-predikat dan mangandungi gabungan leksikal mengikut rumus *Context-free grammar* (CFG). Walaupun tidak dinyatakan penglibatan komponen dengan jelas, namun dapat diringkaskan bahawa komponen yang terlibat mengikut paparan output yang dibuat adalah **kelas kata**, **frasa**, **rumus** dan **penghurai**. Oleh yang demikian, penggunaan komponen ini dijadikan panduan dalam penentuan komponen kajian ini.

2.4.2.2 Penghurai phpSyntaxTree

Penghurai phpSyntaxTree membenarkan pengguna untuk menghasilkan pohon sintaksis secara grafik daripada simbol frasa berbentuk kurungan secara atas talian.

Penghurai ini telah diperkenalkan pada tahun 2003 untuk ayat BI. Walaupun penghurai ini membuat paparan berbentuk huraian hierarki, namun kegunaannya yang memerlukan input berbentuk simbol braket boleh memberi kesukaran kepada pengguna (Rajah 2.7). Walaupun begitu, paparan yang dilakukan boleh menyumbang idea paparan VPS kajian ini semasa fasa pembangunan prototaip dijalankan.



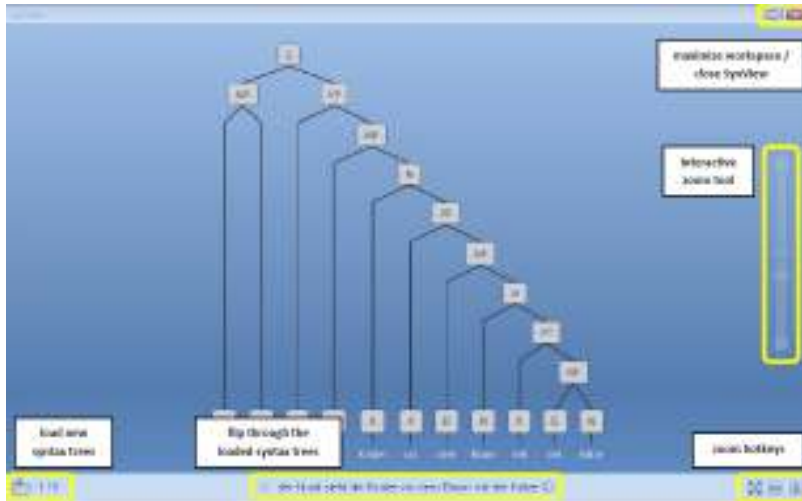
Rajah 2.7. Penghurai **phpSintakTree**

(Sumber: Eisenbach & Eisenbach, 2003)

2.4.2.3 Penghurai SynView

Pada tahun 2009, penghurai SynView direka bentuk oleh sekumpulan pelajar dari Universiti Ruhr Jerman untuk ayat BI. Penghurai ini memerlukan perisian LaTeX dan penganalisis luaran. Input yang diterima berbentuk braket struktur frasa menggunakan notepad. Setelah input dimasukkan, pohon sintaksis dihasilkan menggunakan kaedah bawah-atas dengan membuat pembahagian struktur frasa berdasarkan format input yang diterima. Penghurai ini mempunyai keunikan tersendiri dengan memuatkan teknik pembesar (*zoom*), namun bagi pengguna yang tidak memahami format penulisan input yang diperlukan akan merasa sukar untuk menggunakannya (Rajah 2.8). Walaupun

begitu, kaedah paparan output boleh menyumbang kepada kaedah paparan VPS yang diperlukan dalam kajian ini.



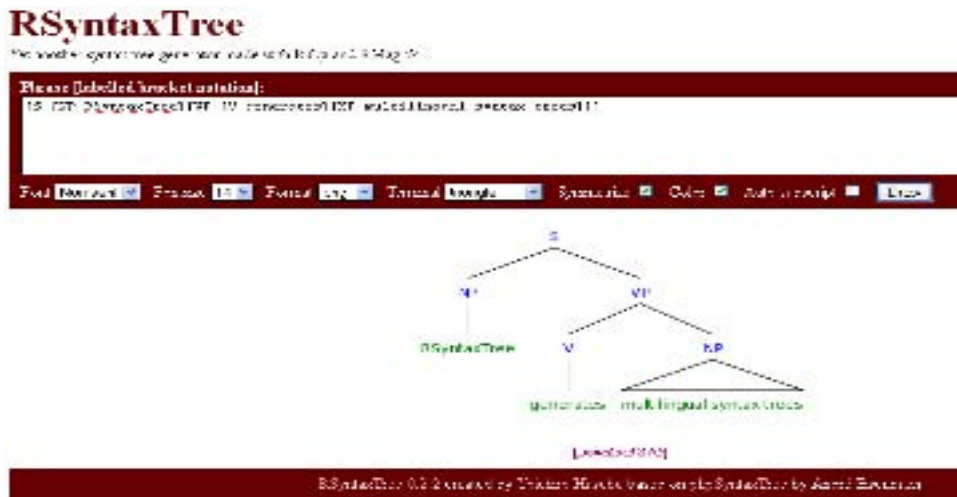
Rajah 2.8. Penghurai SynView

(Sumber: Behrenberg, 2009)

2.4.2.4 Penghurai RSyntaxTree

Penghurai RSyntaxTree direka untuk ayat BI, Jepun, China dan Korea. Penghurai ini menerima input berbentuk simbol kurungan atau bracket struktur frasa. Huraian sintak yang dilakukan tidak memerlukan semakan atau pepadanan pangkalan data. Ini kerana format penerimaan input yang diperlukan akan terus menghasilkan pohon sintaksis.

Walaupun penggunaan akan memberi kesukaran kepada pengguna yang tidak memahami struktur penulisan input yang diperlukan, namun paparan output yang dihasilkan dapat menyumbang idea bagi paparan output pohon sintaksis secara hierarki seperti contoh Rajah 2.9.



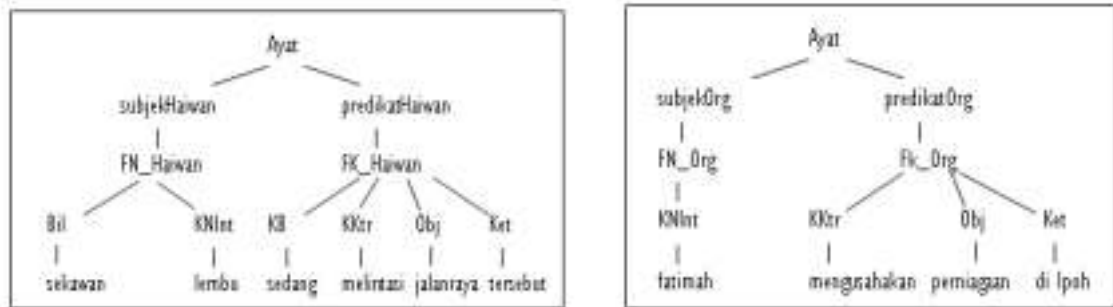
Rajah 2.9. Penghurai RSyntaxTree

(Sumber: Yoichiro, 2012)

2.4.2.5 Penghurai Ayat Bahasa Melayu

Penghurai ayat Bahasa Melayu adalah sebuah kajian yang dilakukan oleh sekumpulan penyelidik dari Universiti Teknologi Petronas yang diketuai oleh Ahmad Izuddin Zainal Abidin (Ahmad Izuddin *et al.*, 2007). Aplikasi yang dibangunkan bertujuan untuk menyemak kebenaran ayat dari segi semantik dan jika ayat yang dimasukkan betul mengikut rumus binaan ayat yang ditetapkan, maka gambaran pohon dipaparkan. Gambaran struktur pembentukan ayat dilakukan bergantung kepada kategori perkataan yang dimiliki dalam ayat yang dimasukkan. Kategori dibahagikan kepada dua jenis iaitu kategori Haiwan dan kategori Manusia. Pembahagian ini dilakukan sebagai langkah untuk membuat penghuraian secara semantik. Ilustrasi pohon yang dilakukan dapat membezakan sama ada ayat tersebut adalah ayat dalam kategori manusia atau haiwan

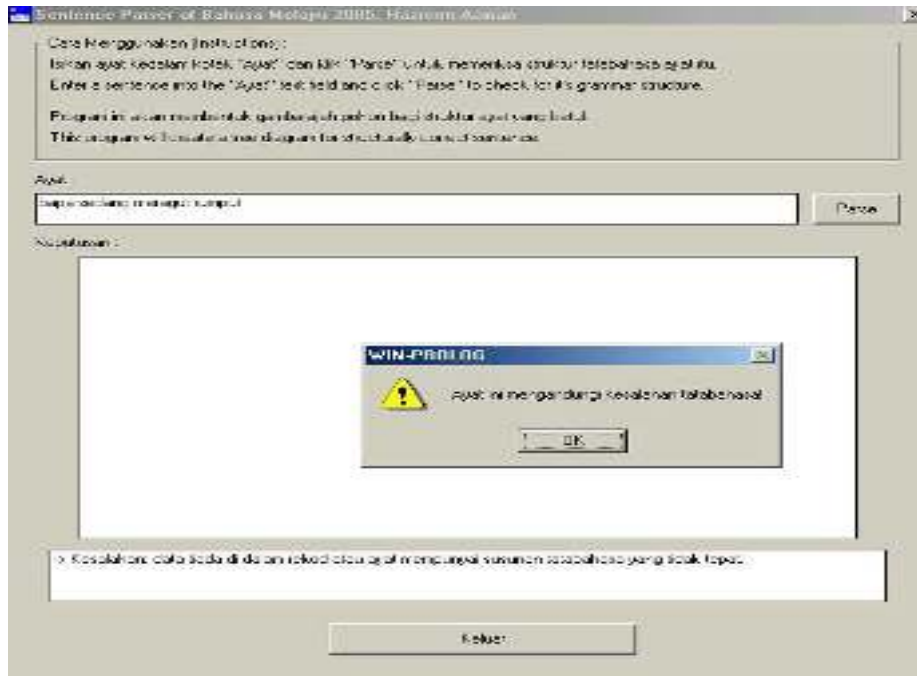
kerana nod subjek dan predikat dibezakan mengikut kategori ini. Rajah 2.10 menunjukkan contoh hasil output yang dikeluarkan jika ayat dimasukkan betul.



Rajah 2.10. Contoh output Penghurai ayat Bahasa Melayu

(Sumber: Ahmad Izuddin *et al.*, 2007)

Jika ayat yang dimasukkan salah mengikut kategori semantik, mesej ralat dipaparkan. Ini bermaksud aplikasi yang dibangunkan akan terlebih dahulu membuat pencarian dan padanan setiap perkataan. Padanan setiap perkataan yang terlibat mestilah dalam kategori yang sama iaitu sama ada Haiwan atau pun Manusia. Sebagai contoh, ayat “bapa saya meragut rumput” akan mengeluarkan mesej ralat kerana “bapa” dikategorikan sebagai Manusia, manakala “meragut” dikategorikan sebagai Haiwan. Mesej ralat dikeluarkan dengan menyatakan bahawa ayat yang dimasukkan mengandungi kesalahan tatabahasa seperti yang ditunjukkan dalam Rajah 2.11.



Rajah 2.11. Antara muka sistem penghurai ayat Bahasa Melayu

(Sumber: Ahmad Izuddin *et al.*, 2007)

Kajian ini menggunakan pendekatan penghuraian secara atas-bawah. Kaedah yang digunakan adalah dengan **menerima dan membaca input, membuat semakan dan padanan**, dan menghasilkan **output** gambaran pohon untuk ayat yang disemak betul atau mesej ralat bagi ayat sebaliknya. Kajian ini berjaya menyemak ayat yang dimasukkan dengan anggaran purata 81% *recall*. Peratusan yang diperolehi adalah berdasarkan kepada bilangan output yang berjaya dikeluarkan. Ini bermaksud 19% ayat tidak berjaya mengeluarkan output.

Walaupun kajian ini telah berjaya menghasilkan output pohon sintaksis, namun output tersebut tidak bercambah untuk membuat output tambahan. Selain itu, pembahagian kelas kata mengikut kategori Haiwan dan Manusia juga memerlukan sebuah pangkalan

data yang besar yang memungkinkan satu perkataan disimpan melebihi sekali untuk setiap kategori. Mesej ralat (Rajah 2.11) yang dikeluarkan juga tidak tepat untuk menyatakan jenis kesalahan yang dikenal pasti. Sebagai contoh, ayat "bapa sedang meragut rumput", ayat ini tidak mempunyai kesalahan tatabahasa sebaliknya kesalahan dari segi semantik. Namun, kaedah penghasilan output yang dilakukan boleh menyumbang idea bagi kegunaan kajian ini.

2.4.2.6 Penghurai Ayat Bahasa Arab

Penghurai ayat bahasa Arab (Shatnawi & Belkhouche, 2012) merupakan sebuah aplikasi untuk menghurai ayat dalam Al-Quran. Kaedah yang digunakan adalah **pembahagian teks kepada ayat, wujudkan fail pengaturcaraan Python untuk setiap ayat, analisis morfologi, padanan kelas kata, padanan rumus, dan output**. Output yang dihasilkan berbentuk pohon sintaksis seperti Rajah 2.12.



Rajah 2.12. Penghurai ayat bahasa Arab

(Sumber: Shatnawi & Belkhouche, 2012)

Kajian tersebut berbeza dengan penghuraian ayat yang lain apabila input yang diterima adalah input berbentuk teks. Teks tersebut dipecahkan kepada ayat sebelum proses penghasilan output dilakukan. Namun kaedah asas penghasilan output pohon sintaksis digunakan dan tidak melibatkan output tambahan selepas pohon sintaksis dipaparkan. Walau bagaimanapun kaedah paparan output dapat menyumbang idea paparan output VPS kajian ini.

2.4.2.7 Penghurai Statistik Ayat Bahasa Melayu

Penghurai statistik untuk mengurangkan keaburan struktur ayat BM telah dibangunkan oleh Noor Hafhizah (2011). Ayat yang dikaji meliputi ayat mudah dengan menggunakan pendekatan penghuraian secara atas-bawah. Sebanyak 1000 ayat dikaji. Ayat yang dikumpul dianalisis untuk mendapatkan rumus binaan ayat untuk menentukan nilai kebarangkalian. Nilai kebarangkalian diperoleh dengan membuat pengiraan pembahagian antara rumus yang dikumpul dibahagikan dengan jumlah ayat terkumpul. Setiap perkataan yang disimpan dalam leksikon akan ditetapkan nilai ini.

Proses yang terlibat dalam membuat penghuraian ayat adalah dengan 1) **menerima dan membaca input**, 2) **mengkelaskan perkataan mengikut kelas kata**, 3) **membuat semakan** dan padanan untuk menentukan kebenaran ayat, 4) mengira nilai jika terdapat perkataan yang kabur iaitu yang melebihi lebih dari satu padanan kelas kata, maka **pengiraan nilai kebarangkalian** dilakukan, dan 5) **output** berbentuk pohon sintaksis dihasilkan. Nilai kebarangkalian dikira dengan menambah kesemua nilai yang dimiliki oleh setiap perkataan, jumlah yang paling tinggi dianggap mempunyai tahap penghuraian yang lebih baik. Nilai yang berbeza adalah pada perkataan yang kabur, di

mana perkataan yang kabur akan mempunyai lebih dari satu nilai bergantung kepada bilangan kelas kata yang dimiliki. Bagi ayat yang tidak sah, mesej ralat dikeluarkan. Output lebih daripada satu pohon sintaksis akan dihasilkan bagi ayat yang kabur dari segi struktur. Selain daripada output pohon sintaksis, nilai kebarangkalian juga dipaparkan. Nilai yang lebih tinggi menunjukkan pohon sintaksis yang lebih baik.

Pohon sintaksis dan nilai kebarangkalian yang dihasilkan dipaparkan secara berasingan berdasarkan paparan output yang ditunjukkan. Output tersebut tidak bercambah dari paparan pohon sintaksis dan kekeliruan juga boleh timbul jika penghasilan output melebihi dari satu kerana nilai kebarangkalian dipaparkan secara berasingan. Walaupun begitu, paparan yang dibuat boleh menyumbang dalam pembentukan paparan output kajian ini dan kaedah yang diikuti boleh dirujuk.

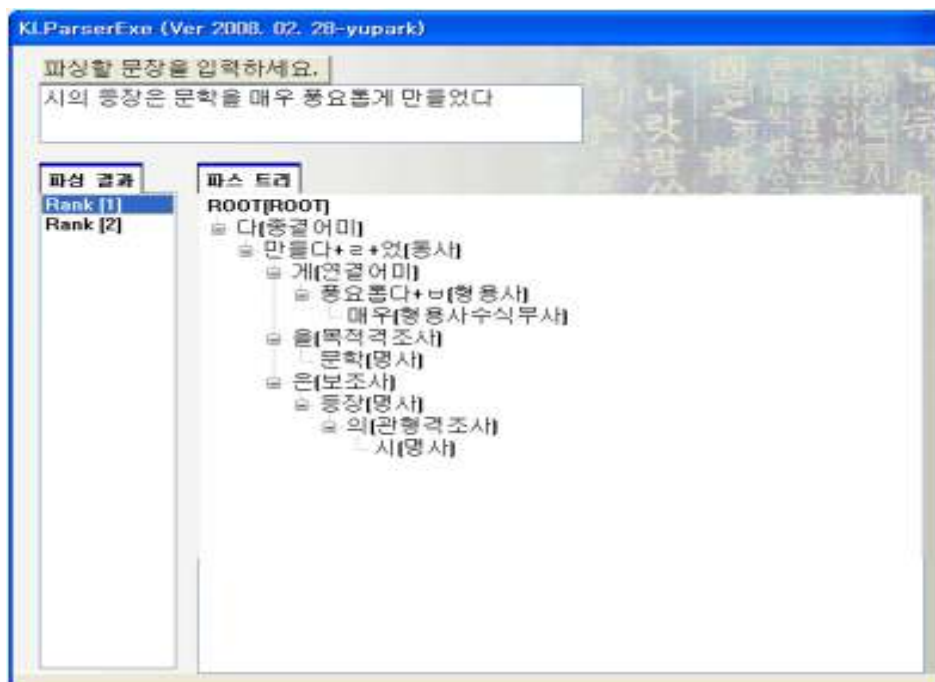
2.4.2.8 Penghurai Statistik Ayat Bahasa Myammar

Penghurai statistik ayat bahasa Myammar (Thant, Htwe & Thein, 2012) menggunakan penandaan kata fungsi. Pohon sintaksis yang dihasilkan adalah untuk kegunaan mesin terjemahan bahasa Myammar ke BI. Penandaan fungsi digunakan untuk menentukan subjek, objek, masa dan lokasi untuk setiap perkataan. Bagi ayat yang salah dikenal pasti berdasarkan rumus, mesej ralat dikeluarkan. Kaedah yang digunakan adalah **menerima input, menanda kelas kata dan kategori, memilih penandaan fungsi bagi setiap kelas kata, memaparkan ayat dengan penandaan fungsi, menghurai penandaan fungsi dengan rumus, dan output**. Aplikasi yang dibangunkan mendapat peratusan *recall* 93% untuk ayat mudah tidak melebihi 15 patah perkataan.

Kelebihan kajian tersebut berbanding kajian lain adalah kaedah penyimpanan maklumat yang memberikan masa, lokasi dan dibahagikan kepada subjek dan objek. Penandaan ini lebih terperinci dan boleh mengurangkan padanan rumus yang salah. Walau bagaimanapun, penghasilan output masih tidak bercambah daripada output pohon sintaksis sebagai output akhir. Mesej ralat bagi ayat salah yang dipadankan, jika ditambah keupayaan rumus untuk membuat cadangan ayat yang betul akan menjadikan aplikasi ini lebih mesra pengguna seperti yang dicadangkan dalam kajian ini.

2.4.2.9 Penghurai Statistik Ayat Korea

Penghurai statistik ayat Korea (Park & Kwon, 2008) dihasilkan untuk mengurangkan kekaburan struktur ayat. Kajian ini menggunakan rumus kebergantungan dan pembahagian ayat. Ayat yang dibahagikan akan dihasilkan pohon sintaksis bagi setiap pembahagian dan penggabungan akan dilakukan pada akhir proses. Kaedah yang digunakan adalah 1) **membina senarai morfem setelah analisis leksikal**, 2) **membuat pembahagian**, 3) **membuang pohon sintaksis yang tidak lengkap untuk setiap pembahagian**, 4) **membuat penggabungan pohon sintaksis**, dan 5) **output**. Hasil uji kaji ke atas 70 ayat mudah daripada buku teks sekolah menengah diperolehi 97% *recall* dan 75% *precision*. Contoh output adalah seperti Rajah 2.13.



Rajah 2.13. Output penghurai statistik bahasa Korea

(Sumber: Park & Kwon, 2008)

Kaedah penghasilan output seakan sama dengan kaedah yang digunakan oleh SSTC, namun tujuan penghasilan berbeza dan output lebih mencerminkan pembentukan ayat berbanding SSTC. Walaupun kajian tersebut tidak bercambah daripada penghuraian pembentukan ayat seperti kajian lain, namun kaedah asas penghuraian ayat digunakan. Kaedah ini boleh menyumbang idea dalam kajian ini.

2.4.2.10 Penghurai Tatabahasa Link Grammar

Link Grammar (LG) adalah sebuah aplikasi yang menghasilkan pohon sintaksis berbentuk kebergantungan (Sleator & Temperley, 1991). Aplikasi ini menghurai ayat dengan cara menerangkan perkaitan antara tatabahasa dalam ayat. Output yang dihasilkan adalah berbentuk separa. Namun, aplikasi ini boleh dijadikan rujukan kerana

pemprosesan ayat yang dilakukan adalah kaedah asas pemprosesan bahasa tabii. LG mempunyai persamaan proses yang dilakukan dalam reka bentuk penghurai ayat BM iaitu, **membaca setiap perkataan, membuat padanan kelas kata dengan rumus, dan menghasilkan output.**

Aplikasi ini tidak melibatkan kaedah semakan ayat. Output akan terus dihasilkan selepas ayat input dimasukkan. Tujuannya untuk menghuraikan perkaitan tatabahasa tanpa output pohon sintaksis secara hierarki. Perbezaan aplikasi penghuraian kebergantungan dengan penghuraian ayat yang lain adalah bahagian semakan. Semakan ayat memerlukan padanan rumus dilakukan sehingga ayat sepenuhnya dapat dipadankan. Jika terdapat salah satu perkataan tidak mempunyai padanan, semakan tidak diteruskan. Walaupun begitu, kaedah pemprosesan yang dilakukan adalah sama dengan kaedah penghuraian ayat yang lain. Kaedah ini menyumbang idea dalam kajian ini sebagaimana yang dijelaskan dalam bahagian seterusnya.

2.4.2.11 Implikasi daripada Penghurai Ayat

Secara keseluruhan, penghurai ayat terdahulu mempunyai kaedah yang seakan sama iaitu membuat semakan, membuat padanan rumus atau kelas kata, dan menghasilkan output. Jadual 2.3 berikut memberi ringkasan kaedah bagi setiap kajian penghurai ayat terlibat.

Jadual 2.3

Ringkasan Sorotan Karya Tentang Algoritma atau Kaedah

Bahasa	Rujukan	Kaedah
BM	Noor Hafhizah (2011)	<ol style="list-style-type: none"> 1. Menerima dan membaca input 2. Setiap perkataan diberikan kelas kata 3. Semakan RSF 4. Pengiraan nilai kebarangkalian 5. Output
	Ahmad Izuddin <i>et al.</i> (2007)	<ol style="list-style-type: none"> 1. Menerima dan membaca input 2. Semakan dan padanan 3. Output
Arab	Shatnawi dan Belkhouche (2012)	<ol style="list-style-type: none"> 1. Pembahagian teks kepada ayat 2. Wujudkan fail pengaturcaraan <code>python</code> untuk setiap ayat 3. Analisis morfologi 4. Padanan kelas kata 5. Padanan rumus 6. Output
Myanmar	Thant, Htwe, dan Thein (2012)	<ol style="list-style-type: none"> 1. Terima input 2. Penandaan kelas kata dan kategori 3. Pemilihan penandaan fungsi bagi setiap kelas kata 4. Memaparkan ayat dengan penandaan fungsi 5. Menghurai penandaan fungsi dengan rumus 6. Output
Korea	Park dan Kwan (2008)	<ol style="list-style-type: none"> 1. Membina senarai morfem setelah analisis leksikal 2. Membuat pembahagian 3. Membuang pohon sintaksis yang tidak lengkap 4. Penggabungan pohon sintaksis 5. Output
BI	Sleator dan Temperley (1991;1993)	<ol style="list-style-type: none"> 1. Membaca setiap perkataan 2. Padanan kelas kata dengan rumus 3. Output

Jadual 2.3 menunjukkan penghasilan output selepas pepadanan dengan rumus adalah proses terakhir yang dilakukan iaitu output berbentuk pohon sintaksis atau pembahagian subjek-predikat. Tiada proses lain yang terlibat seperti yang dilakukan dalam kajian ini iaitu 1) semakan dengan cadangan pembedaan ayat, 2) atribut perkataan, dan 3) VPS melalui ayat contoh. Kajian tersebut menggunakan kaedah asas pemprosesan bahasa tabii iaitu **membahagikan ayat kepada perkataan** atau terus membuat padanan, **padanan dengan rumus** atau kelas kata dan seterusnya menghasilkan **output**. Kaedah asas tersebut dijadikan panduan dalam reka bentuk algoritma VPS dengan output tambahan.

Kajian Ahmad Izuddin *et al.* (2007) tidak melibatkan pembangunan model atau algoritma. Ini telah disahkan oleh Dr. Yong Suet Peng iaitu sebagai salah seorang pengkaji yang terlibat dalam kajian tersebut, melalui perbualan talifon bersama beliau pada 31 Oktober 2014.

"Kajian ini hanya melibatkan *rules* (rumus) sahaja, untuk menilai, prototaip dibangunkan. Tiada model atau algoritma dihasilkan. Hanya satu sahaja penerbitan yang dibuat dan tiada sebarang *report* (laporan) diterbitkan" (Yong Suet Peng, komunikasi peribadi, Oktober 31, 2014).

2.4.3 Penyemak Ayat

Terdapat tiga kajian penyemak ayat terlibat iaitu kajian Rosmah (1995), Suzaimah (2002) dan Tayal, Raghuwanshi dan Malik (2014). Analisis dijalankan ke atas kajian tersebut bertujuan untuk mendapatkan kaedah semakan ayat yang dilakukan. Kaedah ini

akan digunakan dalam reka bentuk algoritma semakan dengan cadangan ayat. Setiap kajian tersebut dijelaskan dalam bahagian seterusnya.

2.4.3.1 Penyemak Sintak Bahasa Melayu

Penyemak sintak yang dibangunkan oleh Rosmah (1995) memberi tumpuan kepada ayat majmuk. Aplikasi yang dibangunkan adalah sebagai alat bantuan kepada sistem terjemahan yang dijalankan oleh Universiti Kebangsaan Malaysia (UKM). Kaedah yang dilakukan adalah **membahagikan ayat kepada perkataan dan kelas kata, membuat padanan rumus, dan output**. Proses bacaan, semakan dan analisis setiap perkataan dalam ayat dilakukan secara huraian atas-bawah. Jika ayat yang dimasukkan menepati rumus yang disediakan, output pembahagian subjek-predikat akan dipaparkan seperti berikut.

Contoh input: Cakera liut itu murah

Output: Subjek: NP (Cakera liut, itu)

Predikat: A (murah)

Output kajian tersebut hanya membahagikan subjek-predikat tanpa paparan jenis kelas kata yang lebih terperinci. Walau bagaimanapun, paparan yang dibuat disemak mengikut padanan rumus. Padanan ini boleh membantu kajian ini dalam mendapatkan gambaran semakan ayat yang dicadangkan.

2.4.3.2 Penyemak Ayat Bahasa Melayu

Suzaimah (2002) menjalankan kajian tentang penyemak ayat BM. Kajian beliau memberi fokus kepada ayat tunggal dan majmuk. Analisis ayat dilakukan dengan

membuat padanan input dengan rumus binaan ayat secara selari menggunakan pengaturcaraan Parlog. Kaedah yang dilakukan adalah 1) **membuat padanan setiap perkataan dengan leksikon**. Setiap rentetan input telah disambungkan kepada klasifikasi utama dan sub-klasifikasi yang ditetapkan dalam leksikon, 2) Jika input tidak memenuhi **padanan rumus**, mesej ralat akan dikeluarkan. Input juga mesti ditulis dalam bentuk kod Parlog yang memerlukan ayat dipisahkan dengan tanda koma dan kurungan, dan 3) menghasilkan **output**.

Contoh, input [ali, makan] akan menghasilkan output berikut:

*P=ayat(frasa_nama(ung_nama(ali)),frasa_kerja(ung_kerja(kata_kerja(makan)))
succeeded.*

Kaedah semakan yang dibuat adalah sama dengan kaedah semakan kajian lain. Kaedah asas pemprosesan bahasa tabii dilakukan. Bezanya output yang dipaparkan berbentuk rumus dan boleh memberi kesukaran pemahaman kepada pengguna bagi membezakan penggunaan simbol bracket. Berbeza juga dengan kajian ini apabila paparan output yang diambil kira adalah paparan VPS bagi struktur pembentukan ayat mengikut teori graf, gestart dan kognitif. Disebabkan pemahaman kognitif pengguna diambil kira, maka paparan output yang jelas tentang pembentukan setiap perkataan, frasa dan kelas kata akan dipaparkan berbanding dengan hanya menggunakan rumus dan simbol bracket.

2.4.3.3 Penyemak Sintak Ayat BI

Penyemak sintak ayat BI (Tayal, Raghuwanshi & Malik, 2014) adalah sebuah penghurai untuk menyemak kebenaran ayat yang dimasukkan. Output yang dihasilkan tidak berbentuk pohon sintaksis tetapi berbentuk ayat sama ada ayat dimasukkan oleh pengguna dikategorikan betul atau pun tidak. Ketepatan output yang diperolehi adalah 81%. Contoh output yang diberikan adalah "*Sentence is syntactically correct* atau *Sentence is syntactically incorrect*". Algoritma yang digunakan mempunyai turutan 1) **masukkan ayat**, 2) **mengkategorikan ayat**, 3) **membuat padanan kelas kata**, 4) **membuat padanan rumus**, dan 5) **menghasilkan output**.

Walaupun output yang dihasilkan berupa ayat pengesahan betul atau tidak ayat yang dimasukkan, namun kaedah asas pemprosesan bahasa tabii juga terlibat. Kajian ini turut berbeza dengan kajian lain apabila ayat terlebih dahulu dikategorikan mengikut jenis sebelum ayat tersebut disemak. Bagaimanapun penghasilan output tidak bercambah untuk membuat paparan berlainan yang menjadikan kajian tersebut serupa dengan kajian semakan yang lain.

2.4.3.4 Implikasi daripada Penyemak Ayat

Perbezaan antara penyemak dengan penghurai ayat adalah dalam bentuk output yang dihasilkan. Output pohon sintaksis dihasilkan bagi penghurai ayat, manakala penyemak biasanya menghasilkan output berbentuk mesej. Namun, semakan diperlukan dalam penghuraian ayat dengan kaedah asas seperti menerima input, membuat padanan, dan output. Jadual 2.4 memberi ringkasan kaedah semakan kajian terdahulu.

Jadual 2.4

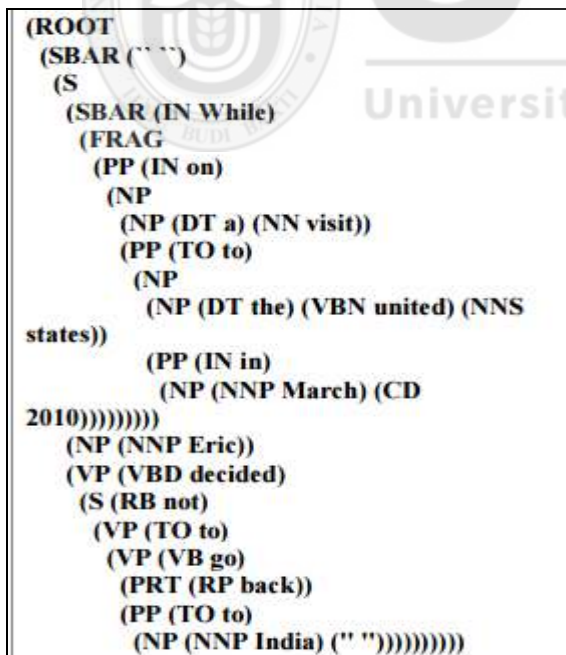
Ringkasan Sorotan Karya Tentang Penyemak Ayat

Bahasa	Rujukan	Kaedah
BM	Rosmah (1995)	1. Membahagikan ayat kepada perkataan dan kelas kata 2. Membuat padanan rumus 3. Output
BM	Suzaimah (2002)	1. Membuat padanan setiap perkataan dengan leksikon 2. Padanan rumus 3. Output
BI	Tayal <i>et al.</i> , (2014)	1. Mengkategorikan ayat 2. Membuat padanan kelas kata 3. Membuat padanan rumus 4. Menghasilkan output

Karya terdahulu seperti dalam Jadual 2.4 menunjukkan kaedah semakan ayat adalah 1) membahagikan/mengkategorikan ayat kepada perkataan, 2) membuat padanan kelas kata/rumus, dan 3) menghasilkan output. Kaedah tersebut adalah kaedah asas pemprosesan bahasa tabii seperti yang turut digunakan dalam kajian penghurai ayat. Justeru itu, dapat dinyatakan bahawa semakan dan penghuraian ayat perlu membuat pembahagian ayat kepada perkataan/kelas kata, membuat padanan rumus, dan menghasilkan output. Kaedah ini akan digunakan dalam penghasilan algoritma VPS dengan output tambahan.

2.4.4 Visualisasi Struktur Ayat: Ekstrak Visualisasi Teks Subjektif

Kajian untuk mengekstrak teks subjektif menggunakan kaedah visualisasi pohon sintaksis telah dijalankan oleh Erfan dan Lili (2014) di Universiti Putra Malaysia. Teks sejarah dipilih sebagai data kajian. Ia bertujuan untuk membantu pemahaman pelajar dalam penerimaan bahagian penting dalam teks digital menggunakan teknik kod warna dan kesan ejaan. Kaedah pembelajaran yang diketengahkan adalah dengan tujuan 1) mengekstrak kelas kata dan bahagian yang diperlukan oleh pengguna, dan 2) visualisasi bahagian yang dipilih dengan kaedah isyarat warna pada perkataan dan pembesaran saiz perkataan, warna dan jarak. Kaedah yang dilakukan adalah **menganalisis leksikal, menetapkan kebergantungan kelas kata**, dan menghasilkan **output**. Contoh output yang diberikan bagi ayat "*While on a visit to the united states in March 2010, Eric decided not to go back to India*" adalah seperti Rajah 2.14.



Rajah 2.14. Output mengekstrak teks subjektif

(Sumber: Erfan & Lili, 2014)

Visualisasi yang dilakukan tidak melibatkan visualisasi pada pohon sintaksis. Namun kaedah pemprosesan ayat yang digunakan boleh dirujuk. Output tersebut juga berbentuk separa dan tidak melibatkan pohon sintaksis secara hierarki. Paparan output yang panjang boleh mengelirukan pengguna kerana penggunaan rumus atau simbol bracket memerlukan penelitian pada output bagi memahami pembentukan ayat yang dibuat.

2.4.5 Cadangan Pembetulan Ayat

Penyemak ayat BM yang dihasilkan oleh Rozana *et al.* (2011) adalah sebuah aplikasi menyemak struktur ayat dengan menterjemah perkataan singkatan dalam ayat input menjadi ayat penuh. Ayat yang diterjemah kemudiannya disemak struktur pembentukan sama ada memenuhi padanan rumus. Bagi ayat yang tidak memenuhi padanan akan diberikan cadangan kesalahan tatabahasa yang sepatutnya terdapat dalam ayat yang dimasukkan. Kaedah yang digunakan adalah, **penterjemahan perkataan singkatan menjadi perkataan penuh, pembahagian perkataan mengikut kelas kata, padanan dengan rumus, dan output.** Contoh output yang diberikan seperti "Ayat ini mengandungi kesalahan, ayat ini harus mengandungi penjodoh_bilangan".

2.4.6 Atribut Perkataan: MALEX

MALEX (MALay LEXicon) adalah sebuah pangkalan data yang dibangunkan untuk menyimpan maklumat linguistik bagi pemprosesan teks BM. Pangkalan data tersebut diperkenalkan oleh Gerry Knowles dan Zuraidah Mohd Don (Chu Min Xian *et al.*, 2016). Kandungan yang terdapat dalam MALEX adalah ejaan, set label, terbitan morfologi dan pangkalan data sebutan (Zuraidah, 2010). Sebanyak 120,000 perkataan

yang diambil daripada novel keluaran DBP diberi kelas kata secara manual (Chu Min Xian *et al.*, 2016).

Pangkalan data yang dibina tidak dihasilkan melalui VPS. Pangkalan data ini juga kurang dibincangkan secara teknikal dari sudut pendekatan penggunaan dan pelaksanaan sistem juga dianggap tidak jelas malah memerlukan usaha yang besar (Chu Min Xian *et al.*, 2016). Namun kaedah penyimpanan kelas kata boleh dirujuk. Kelas kata yang dikumpul secara manual, disimpan dalam pangkalan data dan padanan ke atas kelas kata dibuat melalui padanan struktur sintak. Sebagai contoh, perkataan "keras" mempunyai dua kelas kata iaitu kata kerja dan kata adjektif. Kelas kata ini dikenal pasti melalui padanan dan huraian.

2.4.7 Implikasi Kajian Berkaitan

Jadual 2.5 berikut menyenaraikan 19 kajian berkaitan yang telah dianalisis. Setiap kajian disenaraikan komponen dan kaedah yang digunakan. Hal ini bertujuan untuk mendapatkan jurang kajian yang boleh ditambah baik.

Jadual 2.5 juga menunjukkan kajian penghurai ayat merangkumi komponen perkataan/frasa/ayat, leksikon/pangkalan data, penghurai dan rumus. Komponen tersebut dijadikan asas dalam penghasilan model VPS dengan output tambahan. Manakala kaedah pemprosesan penghurai ayat merangkumi proses pembahagian ayat kepada perkataan, padanan rumus dan menghasilkan output. Oleh itu, kaedah ini dijadikan asas dalam reka bentuk algoritma VPS dengan output tambahan.

Jadual 2.5

Ringkasan Kajian Terdahulu

Kategori	Aplikasi/Kajian	Tahun	Bahasa	Tahap analisis	Komponen	Kaedah
Model	1. SSTC	1998	BI	Sintaksis	Ayat, frasa, pangkalan data, penghurai	Menerima input, padanan, output
	2. Penghurai ayat	2006	BI	Sintaksis	Perwakilan perkataan, perwakilan frasa, struktur ciri-citi, rumus tatabahasa, rumus leksikal	
	3. Penghurai semantik	2008	Matematik	Sintaksis	Penghurai ayat, leksikon, tatabahasa	
Penghurai ayat	4. Lehner's Prolog Tree Drawing	1994	BI	Sintaksis	Simbole Prolog, penghurai	Menerima input, output
	5. phpSintakTree	2003	BI	Sintaksis	Simbol bracket, penghurai	Menerima input, penghurai, output
	6. Penghurai Ayat Bahasa Melayu	2007	BM	Sintaksis dan semantik	Ayat, penghurai	Menerima dan membaca input, semakan, output
	7. SynView	2009	BI	Sintaksis	Simbol bracket, penghurai,	Menerima input, padanan, output
	8. RSyntaxTree	2012	BI, Jepun, China, Korea	Sintaksis	Simbol bracket, penghurai,	Menerima input Penghurai, output

Sambungan Jadual 2.5

Kategori	Aplikasi/Kajian	Tahun	Bahasa	Tahap analisis	Komponen	Kaedah
	9. Penghurai Ayat Bahasa Arab	2012	Arab	Sintaksis	Ayat, penghurai	Pembahagian teks kepada ayat, analisis morfologi, padanan kelas kata dan rumus, output
Penghurai statistik ayat	10. Penghurai statistik ayat Korea	2008	Korea	Sintaksis	Ayat, penghurai	Membina morfem dan analisis leksikal, membuat pembahagian, membuang huraian yang tidak lengkap, menggabungkan huraian, output
	11. Penghurai Statistik Ayat Bahasa Melayu	2011	BM	Sintaksis	Ayat, penyemak, penghurai statistik	Menerima dan membaca input, mengelaskan perkataan mengikut kelas kata, semakan, mengira nilai kebarangkalian, output
	12. Penghurai statistik ayat bahasa Myanmar	2012	Myanmar	Sintaksis		Menerima input, padanan kelas kata, padanan fungsi kelas kata, memaparkan ayat dengan padanan fungsi, padanan rumus, output.
Penghurai tatabahasa	13. Link Grammar	1991	BI	Sintaksis	Ayat, penghurai	Membaca setiap perkataan, padanan pangkalan data dan rumus, output
Penyemak ayat	14. Penyemak sintak	1995	BM	Sintaksis	Ayat, penyemak	Membahagikan ayat kepada perkataan dan kelas kata, padanan rumus, output
	15. Penyemak ayat BM	2002	BM	Sintaksis	Ayat, RSF, leksikon, penyemak	Membuat padanan dengan leksikon, padanan RSF, output

Sambungan Jadual 2.5

Kategori	Aplikasi/Kajian	Tahun	Bahasa	Tahap analisis	Komponen	Kaedah
	16. Penyemak ayat BM	2002	BM	Sintaksis	Ayat, RSF, leksikon, penyemak	Membuat padanan dengan leksikon, padanan RSF, output
	17. Penyemak sintak ayat BI	2014	BI	Sintaksis	Ayat, penyemak	Menerima input, mengkategorikan ayat, padanan kelas kata dan rumus, output
Semakan dengan cadangan ayat	18. Penyemak ayat BM	2011	BM	Sintaksis	Ayat, rumus, penyemak	Terjemah perkataan singkatan kepada perkataan penuh, pembahagian perkataan mengikut kelas kata, padanan rumus, output
Atribut perkataan	19. MALEX	2010	BM	Perkataan	Pangkalan data. teks	Padanan perkataan dengan pangkalan data, output

2.5 Penentuan Model dan Pembuktian Konsep

Komponen asas daripada kajian terdahulu dijadikan panduan dalam reka bentuk model VPS dengan output tambahan. Komponen model ini perlu ditentukan sebelum boleh digunakan dalam reka bentuk algoritma bagi kegunaan fasa pembangunan prototaip. Prototaip tersebut digunakan untuk membuat pembuktian konsep. Kaedah yang digunakan dalam penentuan model dan pembuktian konsep dijelaskan dalam bahagian seterusnya.

2.5.1 Penentuan Model Kajian Terdahulu

Penilaian pakar (*expert review*) adalah salah satu cara penentuan model (Pathiah, 2012) seperti yang dilakukan dalam penentuan model kajian Syamsul Bahrin (2011) dan Muhamad Shahbani (2012). Proses dalam penentuan model adalah pemilihan pakar bidang dan pembinaan instrumen penilaian.

Syamsul Bahrin (2011) telah melibatkan empat orang pakar bidang dengan menggunakan lapan dimensi soalan yang merangkumi keberkesanan penggunaan komponen model. Terdapat empat kaedah penentuan model yang digunakan iaitu 1) penerangan mengenai tujuan penentuan, 2) peluang untuk pakar menjalankan penilaian, 3) analisis dapatan, dan 4) membuat penambahbaikan model. Model akhir setelah penentuan digunakan untuk membangunkan sebuah prototaip bagi menjalankan pembuktian konsep.

Muhamad Shahbani (2012) menggunakan kaedah penentusahan yang sama dengan Syamsul Bahrin (2011) iaitu melibatkan empat kaedah melalui enam orang pakar penilai. Setelah model ditentukan, pakar penilai diminta membuat rumusan ke atas penilaian secara keseluruhan. Hasil dapatan digunakan untuk membuat penambahbaikan model untuk membina prototap kajian bagi menjalankan pembuktian konsep.

Walaupun penentusahan Muhamad Shahbani (2012) melibatkan kaedah rumusan pada akhir proses, namun kaedah ini sebenarnya telah termasuk dalam penilaian yang dilakukan pakar penilai. Sebagaimana yang dilakukan dalam kajian Syamsul Bahrin (2011) iaitu memberi peluang kepada pakar penilai membuat penilaian setelah penerangan dibuat. Rumusan boleh disertakan dalam instrumen penilaian.

Oleh itu, berdasarkan kaedah penentusahan kajian terdahulu, maka empat kaedah penentusahan diambil kira iaitu 1) penerangan mengenai tujuan penentusahan, 2) pakar menjalankan penilaian, 3) analisis dapatan, dan 4) membuat penambahbaikan model. Model ini akan digunakan untuk mereka bentuk algoritma VPS dengan output tambahan. Seterusnya prototap kajian dibina untuk menjalankan pembuktian konsep seperti yang dilakukan dalam kajian terdahulu yang dijelaskan dalam bahagian seterusnya.

2.5.2 Pembuktian Konsep Kajian Terdahulu

Bagi menjawab persoalan kajian ketiga iaitu "Bagaimana cara untuk memastikan model dan algoritma yang dihasilkan adalah tepat?", maka penilaian hasil kajian perlu

dilakukan. Penilaian ini adalah sebagai pembuktian konsep dalam model VPS yang diterjemahkan dalam bentuk turutan algoritma VPS dengan output tambahan.

Noor Hafhizah (2011) dan Ahmad Izuddin *et al.* (2007) menjelaskan dapatan kajian menggunakan metrik penilaian Parseval. Metrik ini adalah untuk mendapatkan jumlah peratusan output pohon sintaksis mengikut formula *recall* atau *precision*. Walaupun kaedah penilaian dapatan tidak dinyatakan secara terperinci dalam kedua-dua kajian tersebut, namun pemilihan metrik penilaian ini dapat memberi idea dalam pemilihan penilaian yang perlu dibuat.

Noor Hafhizah (2011) menggunakan sebuah prototaip untuk mendapatkan hasil ketepatan output pohon sintaksis. Kaedah yang digunakan adalah 1) pengumpulan ayat uji kaji, 2) uji kaji melalui prototaip, dan 3) analisis dapatan output. Dapatan output dinilai menggunakan metrik *recall*, *precision* dan *f-score* berdasarkan pola ayat dasar BM.

Ahmad Izuddin *et al.* (2007) juga menjalankan uji kaji prototaip melibatkan tiga kaedah iaitu 1) pengumpulan ayat, 2) uji kaji, dan 3) analisis dapatan. Pengumpulan ayat yang dilakukan adalah melibatkan guru sekolah rendah dan sekolah menengah untuk menyediakan ayat mudah mengikut skop kajian. Ayat yang dikumpul dimasukkan ke dalam prototaip untuk diuji. Seterusnya dapatan output dinilai menggunakan metrik penilaian *recall*.

Oleh yang demikian, kaedah pembuktian konsep melibatkan tiga kaedah iaitu pengumpulan ayat yang boleh diambil dari data kajian, uji kaji prototaip dan menganalisis output. Output ini mestilah dianalisis menggunakan metrik penilaian Parseval sebagai kaedah menilai output pohon sintaksis.

2.6 Jurang Kajian

Kajian terdahulu telah dianalisis untuk mendapatkan komponen dan kaedah dalam penghurai ayat. Komponen perkataan/frasa/ayat, leksikon/pangkalan data, penghurai ayat dan rumus adalah komponen terlibat dalam sebuah model penghurai ayat. Kaedah asas pemprosesan bahasa tabii pula terlibat dalam membuat penghuraian ayat iaitu membahagikan ayat kepada perkataan, padanan dengan rumus atau kelas kata dan menghasilkan output.

Kajian terdahulu ditambah baik dengan menambah komponen 1) cadangan pembetulan ayat, 2) atribut perkataan, dan 3) VPS melalui ayat contoh. Namun adakah komponen tambahan tersebut sememangnya penting dan tidak pernah dikaji? untuk mendapatkan jawapan ini, persoalan berikut perlu dijawab. "Apakah jurang yang perlu diterokai daripada kajian penghurai ayat? Oleh yang demikian, pemahaman ke atas jurang berikut perlu diambil kira.

- I. Mengapa cadangan pembetulan perlu disertakan sedangkan Jadual 2.2 telah menyatakan terdapat penyemak dengan cadangan ayat telah diperkenalkan (item 18)?

- II. Mengapa atribut perkataan perlu disertakan sedangkan MALEX telah menghasilkan korpus perkataan (item 19)?
- III. Terdapat 19 kajian yang telah diperkenalkan sejak 1980-an, masih tiadakah VPS melalui ayat contoh diketengahkan?
- IV. Apakah komponen lain yang diperlukan?
- V. Mengapa tahap pemprosesan sintaksis dipilih sedangkan terdapat kajian yang menganalisis semantik?
- VI. Mengapa jurang dalam I-IV penting? dan mengapa kajian perlu dilakukan?

Jurang I

Cadangan pembedaan yang dicadangkan dalam Rozana *et al.* (2011) adalah cadangan berbentuk kesalahan tatabahasa. Namun, cadangan pembedaan bagi ayat yang tidak tepat mengikut rumus binaan ayat BM tidak dilakukan. Sebagai contoh, kesalahan tatabahasa akan dipaparkan seperti "ayat harus mengandungi penjodoh bilangan" kerana ayat yang dimasukkan dianggap tidak lengkap tanpa memberi cadangan ayat yang lain.

Jurang II

MALEX yang dapat dirujuk dalam Zuraidah (2010) adalah sebuah pangkalan data yang mengandungi maklumat linguistik. Antara kandungan yang terdapat dalam MALEX mempunyai persamaan dengan kandungan atribut perkataan kajian ini dari segi penggunaan kelas kata. Walau bagaimanapun, maklumat yang disimpan dalam MALEX tidak dipaparkan hasil pautan pada VPS. Ia juga adalah sebuah pangkalan data. Manakala atribut perkataan yang digabung dalam VPS kajian ini diperolehi hasil pautan bagi satu perkataan yang terdapat dalam sebuah ayat.

Jurang III

Pohon sintaksis yang dihasilkan dalam kajian terdahulu adalah output terakhir. Walaupun terdapat kajian yang menggunakan kaedah visualisasi seperti kajian Erfan dan Lili (2014), namun visualisasi yang dilakukan adalah pada peringkat mengekstrak ayat. Teks yang dipaparkan boleh dibesarkan saiz perkataan dan warna pengecaman pada kelas kata tertentu diberikan. Teknik ini diaplikasikan sebelum ayat diekstrak untuk membuat pohon sintaksis. Perbezaan dengan kajian ini adalah, visualisasi dilakukan setelah paparan pohon sintaksis dipaparkan di mana teknik visualisasi diberikan dengan menyediakan pautan pada setiap nod dalam pohon sintaksis untuk menghasilkan atribut perkataan.

Jurang IV

Semakan ayat memerlukan rumus tatabahasa dikaji. Rumus binaan ayat BM perlu dikaji berdasarkan teori *X-bar* kerana rumus ini adalah peningkatan kepada rumus binaan ayat BM yang universal. Selain itu, rumus binaan ayat juga penting dalam reka bentuk model, algoritma dan prototaip kajian.

Selain itu, model dan algoritma yang mendasari pembangunan VPS berserta output tambahan perlu diperkenalkan. Hal ini kerana model atau algoritma kajian terdahulu tidak bercambah dari output pohon sintaksis dan skop tujuan yang berbeza. Model dan algoritma kajian terdahulu juga tidak diketengahkan secara terperinci dan tidak sepenuhnya merangkumi skop kajian ini. Sebagai contoh, walaupun SSTC menghasilkan model reka bentuk, namun tidak melibatkan kaedah semakan ayat. Ini

dijelaskan oleh Prof Madya Dr. Tang Enya Kong melalui maklum balas e-mel beliau kepada penulis.

"I have not used the SSTC for syntax checking so far as it is more appropriate for machine translation due to the limitation of the Example-based parsing technique which has been implemented (deterministic parsing)" (Tang Enya Kong, komunikasi peribadi, April 4, 2011).

Kajian Ahmad Izuddin *et al.* (2007) juga tidak melibatkan pembangunan model atau algoritma. Ini telah disahkan oleh Dr. Yong Suet Peng iaitu sebagai salah seorang pengkaji yang terlibat dalam kajian tersebut, melalui perbualan talifon bersama beliau pada 31 Oktober 2014 (rujuk muka surat 60).

Secara ringkas, hasil komunikasi peribadi yang dilakukan, VPS ayat penyata BM masih belum mendapat perhatian dalam kalangan pengkaji. Objektif kajian ini untuk menghasilkan model dan algoritma bagi output tambahan sememangnya masih belum diwujudkan. Selain itu, rumus binaan ayat BM yang lengkap untuk ayat penyata masih belum diwujudkan kerana ayat BM adalah bersifat tidak tetap dan boleh berubah mengikut kesesuaian keadaan. Justeru itu, rumus binaan ayat yang lengkap untuk pembangunan VPS ayat penyata BM perlu dikaji dan dianalisis.

Jurang V

Bahagian 1.4.2 Skop struktur ayat telah menjelaskan peri pentingnya tahap pemrosesan sintaksis. Aspek sintaksis dipilih kerana aspek ini adalah aspek utama yang harus difokuskan berbanding aspek lain seperti semantik. Aspek ini merangkumi kajian struktur pembentukan ayat. Selain penambahbaikan yang dilakukan dengan mencadangkan cadangan pembedulan ayat, atribut perkataan dan VPS ayat contoh yang pertama kali diperkenalkan, maka aspek sintaksis terlebih dahulu perlu diterokai. Kejayaan aspek ini akan dapat dikembangkan kepada aspek lain seperti semantik.

Jurang VI

Komponen cadangan pembedulan ayat perlu disertakan sebagai bantuan pemahaman dalam menghasilkan ayat yang lebih baik. Penghasilan ayat yang gramatis perlu berlandaskan kepada rumus iaitu gabungan nahu yang lengkap agar ayat tersebut diucap dan ditulis dengan betul. Oleh yang demikian, bagi ayat yang dimasukkan tidak menepati rumus pembentukan ayat mengikut standard pembentukan ayat BM, maka cadangan pembedulan ayat akan dapat membantu pengguna.

Atribut perkataan pula sebagai alternatif sampingan dalam memahami sesebuah perkataan. Sesebuah perkataan boleh menghasilkan ayat berbeza. Selain memberi contoh ayat berbeza, imej yang diberikan akan membantu pengguna yang kurang mahir berkenaan sesebuah perkataan.

Sebuah ayat akan mengandungi beberapa perkataan. Setiap perkataan yang terkandung pula boleh menghasilkan ayat berbeza. Percambahan ini yang hendak diketengahkan agar pembentukan sesebuah ayat bukan hanya dipelajari pembentukan rumus terlibat, malah atribut dan ayat berbeza yang boleh dihasilkan juga dapat membantu pengetahuan yang lebih baik.

Persoalannya, mengapa kajian ini perlu diketengahkan? Oleh kerana kajian terdahulu mementingkan paparan pohon sintaksis dan membuat semakan ayat, maka kajian yang boleh membuat penambahbaikan seperti output tambahan perlu dilakukan. Jadual 2.5 (rujuk muka surat 68) telah menyenaraikan kajian terdahulu termasuk kaedah yang dijalankan. Oleh kerana kajian tersebut hanya menghadkan kepada pohon sintaksis sebagai output akhir, maka sumbangan baharu berupa penambahbaikan output tambahan perlu dikaji. Hal ini supaya pohon sintaksis boleh digunakan dalam konteks yang pelbagai. Sebagai contoh, pohon sintaksis boleh menghasilkan atribut berbeza seperti imej, kelas kata, kata terbitan dan ayat contoh bagi setiap perkataan.

Apakah pentingnya sumbangan yang diajukan? Pembentukan sesebuah ayat boleh difahami dengan kaedah berbeza. Hanya dengan menggunakan sebuah ayat, pembentukan ayat akan dapat menghasilkan atribut berbeza melalui pautan pada setiap perkataan. Selain itu, kajian terdahulu masih ditahap pohon sintaksis sebagai output akhir, sehingga tahun 2014, pembentukan output masih sama ekorannya pemahaman pembentukan ayat hanya melibatkan pembentukan nahu semata. Namun, melalui pembentukan nahu tersebut, komponen lain seperti cadangan ayat dan atribut perkataan

boleh disertakan. Selain itu, model dan algoritma yang mendasari output tambahan ini masih tidak diperkenalkan. Justeru itu, penambahbaikan 1) model VPS berserta dengan output tambahan 2) algoritma VPS dengan semakan serta cadangan pembedahan ayat, 3) rumus X -bar bagi kegunaan VPS ayat penyata BM, dan 4) pengecaman atribut perkataan menjadi sumbangan baharu untuk dikaji.

2.7 Rumusan Bab Dua

Bab dua memberi penjelasan mengenai latar belakang kajian yang menjelaskan mengenai perkaitan bidang pengkomputeran linguistik dan NLP. Perkaitan ini adalah berdasarkan bidang pengkomputeran linguistik sebagai sub-bidang dalam bidang linguistik umum. Perkaitan tersebut telah merungkai kajian berkenaan pohon sintaksis iaitu kajian yang melibatkan pemahaman mengenai ilmu bahasa dan pengkomputeran. Perkaitan ini juga dikaitkan dengan penggunaan teori terlibat iaitu teori X -bar, teori graf, teori gestalt dan teori beban kognitif.

Kaedah penjelasan bab ini juga adalah berlandaskan kepada persoalan kajian. Bagi menjawab persoalan kajian pertama berkenaan rumus binaan ayat BM, maka kajian berkaitan rumus ini diketengahkan. Penjelasan tersebut telah menemukan bahawa rumus X -bar adalah rumus yang ditambah baik bagi kegunaan dalam pembentukan ayat BM.

Persoalan kajian kedua pula dijelaskan dengan membuat sorotan kritikal terhadap kajian berkaitan model, algoritma dan prototaip dalam bidang penghurai ayat. Penjelasan ini termasuklah kajian mengenai penentusahan model dan pembuktian konsep bagi menjawab persoalan ketiga. Setiap kajian terlibat difahami dan dijelaskan ruang yang

boleh ditambah baik dalam bahagian Jurang Kajian yang menonjolkan sumbangan kajian ini. Bahagian tersebut memberi penekanan kepada model dan algoritma, rumus binaan ayat, VPS, semakan dengan cadangan ayat, dan atribut perkataan.

Hasil analisis yang dilakukan ke atas sorotan karya terlibat, dapat disenaraikan komponen sesebuah penghurai ayat. Komponen ini mengandungi perkataan/frasa/ayat, leksikon, penghurai, dan rumus. Manakala kaedah yang boleh diikuti pula seperti membahagikan ayat kepada perkataan, padanan dengan rumus, dan output.

Setelah analisis kritikal dilakukan, sumbangan yang dicadangkan adalah seperti 1) model VPS berserta dengan output tambahan 2) algoritma VPS dengan semakan serta cadangan pembetulan ayat, 3) rumus $X\text{-bar}$ bagi kegunaan VPS ayat nyata BM, dan 4) pengecaman atribut perkataan. Oleh yang demikian, sumbangan tersebut digunakan untuk menentukan kaedah terlibat dalam Bab Tiga seterusnya.

BAB TIGA

METODOLOGI KAJIAN

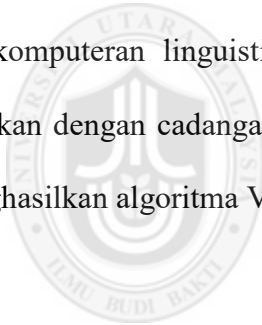
3.0 Pengenalan

Metodologi kajian yang digunakan adalah penyelidikan reka bentuk (PR). Sebarang kajian yang menggunakan PR bermatlamat untuk menghasilkan reka bentuk model, algoritma atau prototaip (Prat, Comyn-Wattiau & Akoka, 2014; Hevner, March & Park, 2004; Norshuhada & Shahizan, 2010; Vaishnavi & Kuechler, 2008). Fasa yang terlibat dalam PR melibatkan kaedah 1) pengetahuan pernyataan masalah, 2) cadangan, 3) pembangunan, 4) penilaian dan 5) rumusan (Muhamad Shahbani, 2012; Vaishnavi dan Kuechler, 2008). Oleh itu, untuk kajian ini, Fasa 1 menggunakan kaedah pengetahuan pernyataan masalah, Fasa 2 mencadangkan kaedah mengkategorikan dan menganalisis struktur ayat untuk mendapatkan rumus serta mencadangkan reka bentuk model dan algoritma, Fasa 3 menggunakan kaedah pembangunan, Fasa 4 menjalankan kaedah penilaian dan Fasa 5 memberi rumusan.

Berdasarkan kepada sumbangan yang dicadangkan hasil analisis kritikal dalam Bab Dua, maka kaedah kajian ini hendaklah berasaskan kepada reka bentuk model dan algoritma. Hal ini kerana, analisis ke atas sorotan karya mendapati model dan algoritma pohon sintaksis dengan output tambahan masih tidak diperkenalkan (rujuk bahagian 2.6). Oleh itu, kajian ini telah memilih metodologi kajian PR kerana 1) matlamat PR adalah untuk reka bentuk model, algoritma atau prototaip dan 2) lima fasa dalam PR adalah sepadan dengan sumbangan dan objektif kajian seperti yang dijelaskan melalui Rajah 3.1.

Fasa pertama adalah untuk mengetahui jurang yang terdapat dalam kajian penghurai ayat. Dalam fasa ini, komponen penghuraian ayat disenaraikan bagi mendapatkan komponen dan kaedah terlibat untuk tujuan penambahbaikan output pohon sintaksis. Kajian penghuraian ayat terlibat dianalisis untuk pengukuhan pernyataan masalah dan skop kajian. Penjelasan telah dibuat dalam Bab Satu dan Bab Dua.

Fasa kedua mengkategorikan dan menganalisis struktur ayat BM bagi mendapatkan rumus binaan ayat. Fasa ini juga untuk mereka bentuk model dan algoritma VPS dengan output tambahan. Model atribut perkataan direka bentuk terlebih dahulu untuk mendapatkan model VPS. Model ini ditentukan oleh pakar BM dan pakar bidang pengkomputeran linguistik. Setelah model VPS dipersetujui, reka bentuk algoritma semakan dengan cadangan pembetulan ayat dihasilkan. Algoritma ini diteruskan untuk menghasilkan algoritma VPS dengan output tambahan.



UUM
Universiti Utara Malaysia

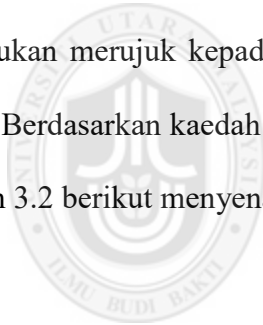
Fasa 1: Pengetahuan pernyataan masalah

SUMBANGAN	OBJEKTIF	
<ol style="list-style-type: none"> 1. Rumus $X\text{-bar}$ bagi kegunaan VPS 2. Pengecaman atribut perkataan 3. Model VPS berserta output tambahan 4. Algoritma VPS dengan semakan serta cadangan pembedulan ayat 	<ol style="list-style-type: none"> 1. Mengkategorikan ayat BM untuk mendapatkan rumus yang tepat 2. Membina model, algoritma dan prototaip kajian 	<p>Fasa 2: Mengkategorikan dan menganalisis struktur ayat untuk mendapatkan rumus, reka bentuk model dan algoritma VPS dengan output tambahan</p> <p>Fasa 3: Pembangunan</p>
	<ol style="list-style-type: none"> 3. Menilai output VPS dan cadangan pembedulan ayat 	<p>Fasa 4: Penilaian</p>
		<p>Fasa 5: Rumusan</p>

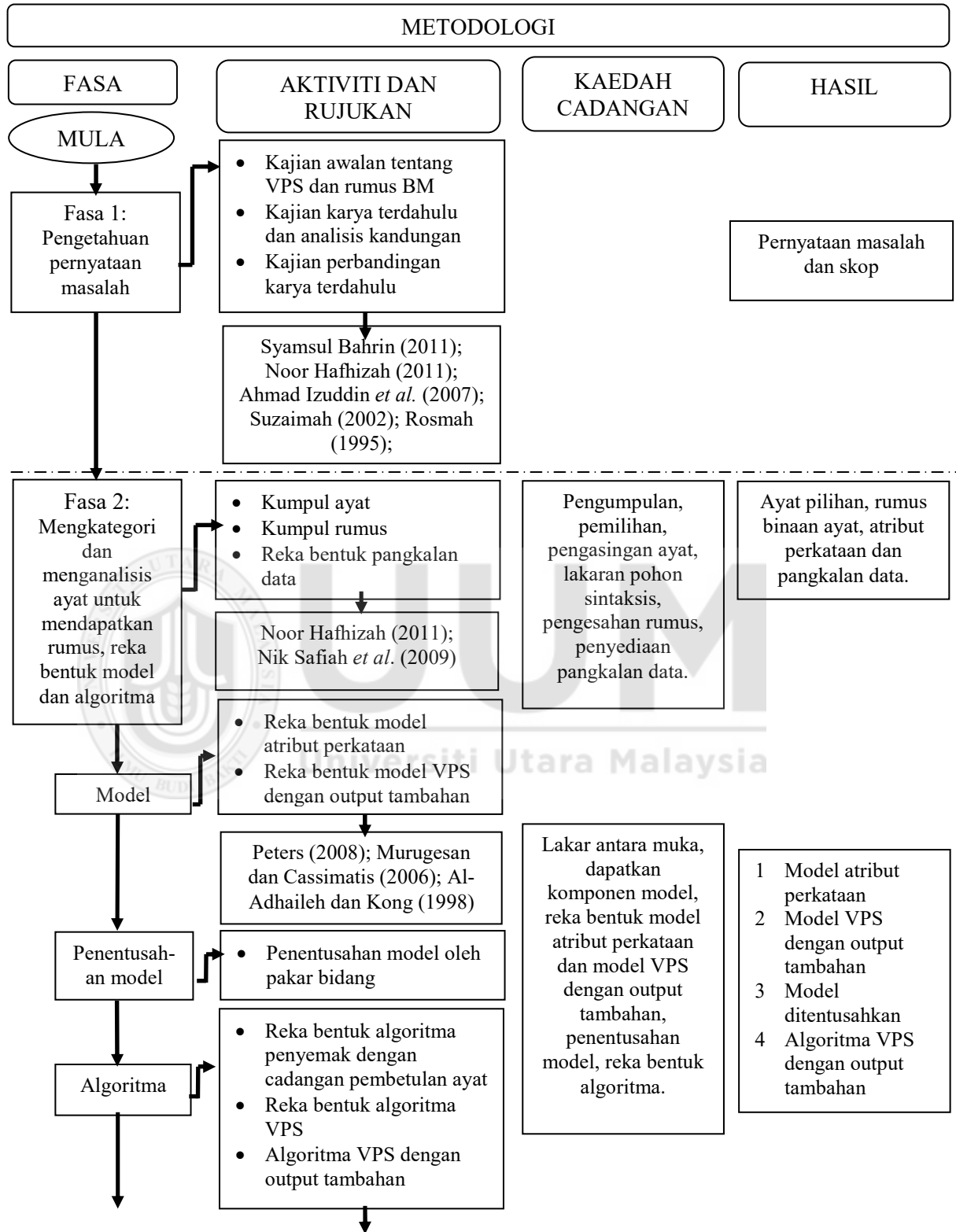
Rajah 3.1. Perkaitan metodologi dengan sumbangan dan objektif kajian

Fasa ketiga adalah bertujuan untuk membina prototaip berdasarkan algoritma VPS dengan output tambahan. Antara muka prototaip terlebih dahulu dilakar. Lakaran ini dihubungkan dengan pangkalan data. Setiap proses dikodkan mengikut turutan algoritma. Seterusnya, fasa terakhir untuk menjalankan penilaian prototaip dilakukan. Penilaian ini dibahagikan mengikut dua kaedah. Kaedah penilaian Parseval dan kaedah penilaian pengguna. Penilaian ini melibatkan pengguna akhir untuk mendapatkan skor min bagi kepuasan subjektif dan kognitif pengguna.

Secara berstruktur, metodologi kajian ini dijelaskan melalui Rajah 3.2. Penjelasan dibuat mengikut pembahagian fasa, aktiviti, kaedah dan hasil yang dicapai. Setiap aktiviti yang dilakukan merujuk kepada kaedah kajian terdahulu seperti yang dijelaskan dalam Bab Dua. Berdasarkan kaedah tersebut, kaedah kajian ini dicadangkan mengikut skop kajian. Rajah 3.2 berikut menyenaraikan setiap aktiviti yang terlibat.

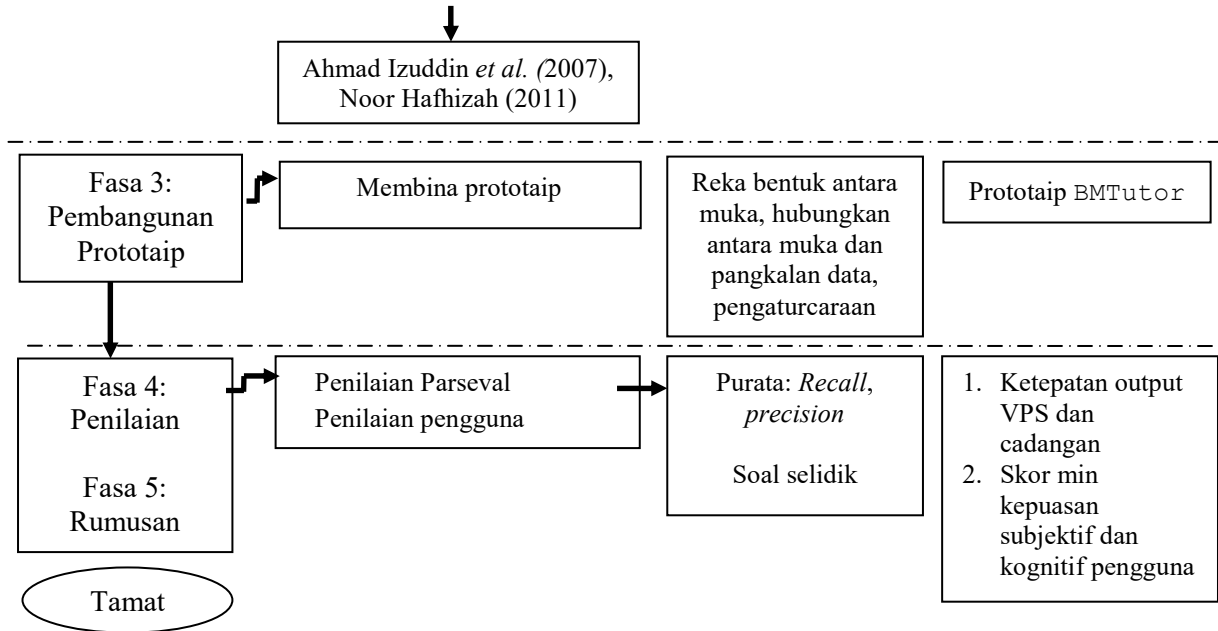


UUM
Universiti Utara Malaysia



Rajah 3.2. Metodologi kajian PR.

Sambungan Rajah 3.2



3.1 Fasa 1: Pengetahuan Pernyataan Masalah

Dalam fasa pertama, pernyataan masalah diperkukuhkan dengan melakukan kajian awalan tentang VPS dan rumus binaan ayat BM. Selain itu, kajian karya terdahulu dilakukan dengan membuat analisis kandungan dan perbandingan bagi mendapatkan jurang yang boleh ditambah baik. Berdasarkan kepada maklumat yang diperoleh, objektif, skop dan persoalan kajian ditetapkan seperti yang telah dijelaskan dalam Bab Satu.

3.1.1 Kajian Awalan

Komunikasi secara peribadi dilakukan dengan pakar bidang iaitu Prof Madya Dr. Tang Enya Kong. Komunikasi melalui panggilan telefon dilakukan untuk mengetahui tentang kajian pemrosesan ayat yang telah dan sedang dilakukan untuk BM. Selain itu, komunikasi melalui e-mel juga turut membantu. Komunikasi peribadi dengan Prof.

Emeritus Datuk Dr. Nik Safiah Karim sebagai pakar BM juga telah dilakukan untuk mengetahui tentang rumus binaan ayat BM yang telah dihasilkan (rujuk muka surat 39).

Hasil komunikasi peribadi, kajian berkenaan VPS ayat BM masih belum mendapat perhatian dalam kalangan pengkaji. Disebabkan perkara tersebut, model dan algoritma output tambahan pohon sintaksis masih belum diwujudkan. Rumus binaan ayat penyata BM untuk kegunaan VPS masih belum diwujudkan kerana ayat BM adalah bersifat tidak tetap dan boleh berubah mengikut kesesuaian keadaan. Justeru itu, rumus binaan ayat ini perlu dikaji dan dianalisis. Analisis ini dimulakan dengan membuat analisis kandungan karya terdahulu.

3.1.2 Kajian Karya Terdahulu dan Analisis Kandungan

Analisis sorotan karya terdahulu dilakukan bertujuan untuk merapatkan jurang dalam kajian pohon sintaksis. Hal ini kerana kajian dalam bidang ini telah dilakukan sejak tahun 1980-an namun paparan output masih tidak bercambah dari penghasilan pohon sintaksis. Kaedah yang digunakan, model atau algoritma yang telah direka bentuk dianalisis bagi pengukuhan pernyataan masalah dan skop kajian. Penjelasan secara jelas telah ditunjukkan dalam Bab Dua.

3.1.3 Kajian Perbandingan Karya Terdahulu

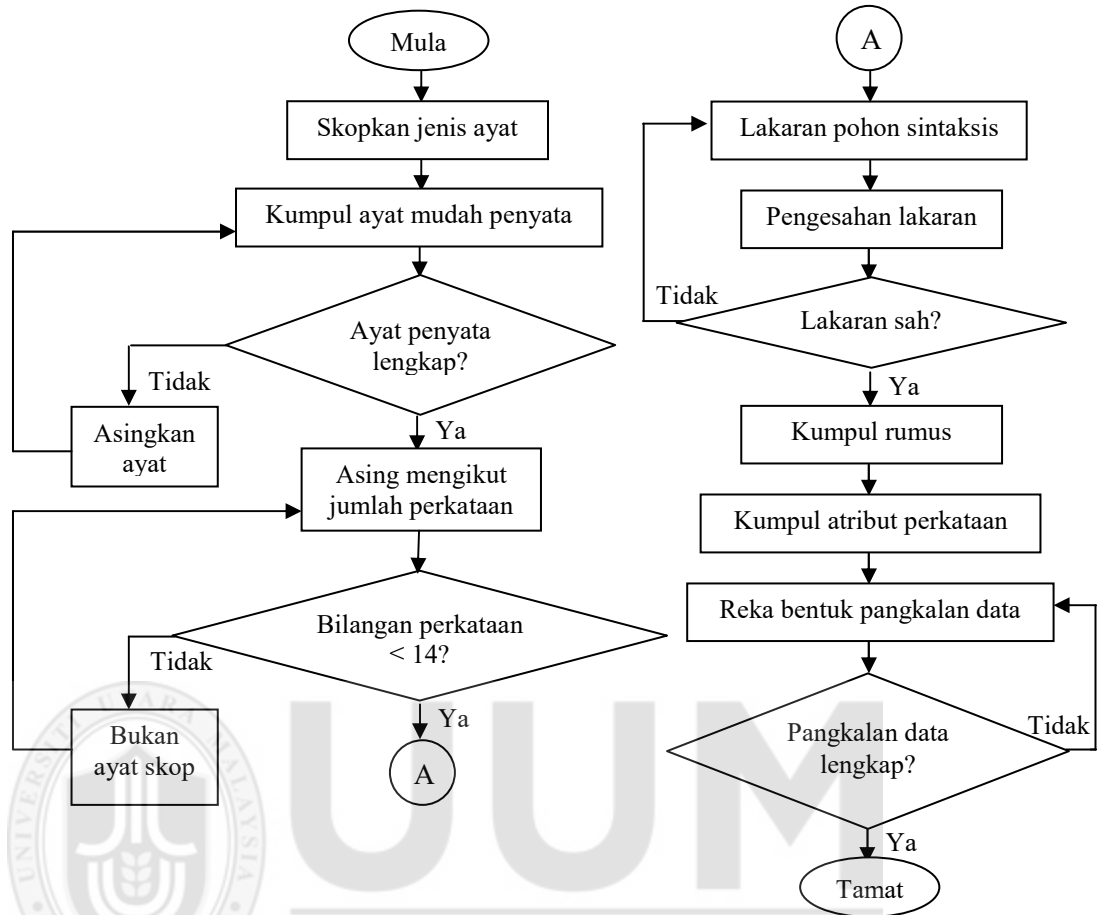
Perbandingan karya terdahulu dilakukan mengikut persoalan kajian. Sebagai contoh, adakah rumus binaan ayat BM telah diperkenalkan untuk VPS? Persoalan tersebut telah membawa kepada penemuan kajian Noor Hafhizah (2011) yang memperkenalkan rumus

asas pembentukan ayat BM. Selain itu, jadual perbandingan ulasan karya terlibat telah dijelaskan dalam Bab Dua (Jadual 2.5).

3.2 Fasa 2: Rumus *X-bar*, Model dan Algoritma

Fasa ini adalah untuk mencapai objektif pertama kajian iaitu "Mengkategorikan dan menganalisis struktur ayat BM untuk mendapatkan rumus yang tepat" dan objektif kedua iaitu "Membina model dan algoritma VPS dengan output tambahan serta prototaip kajian sebagai alat pembuktian konsep".

Fasa mengkategorikan dan menganalisis struktur ayat BM mengikut skop kajian dimulakan dengan sesi pengumpulan, pemilihan, pengasingan, lakaran pohon sintaksis, pengesahan, dan penyediaan pangkalan data ayat penyata BM. Aktiviti ini dibahagikan kepada tiga proses. Proses pertama adalah proses pengumpulan ayat untuk menghasilkan rumus *X-bar*, mendapatkan perkataan dan ayat. Diikuti dengan proses pengesahan rumus oleh pakar BM bagi menentukan rumus yang digunakan adalah tepat. Proses ketiga adalah proses pengumpulan maklumat yang telah dikumpul. Proses yang terlibat dalam fasa ini ditunjukkan dalam Rajah 3.3. Rumus yang diperolehi dalam kajian ini dijelaskan dalam Bab Lima.



Rajah 3.3. Carta alir proses mengkategorikan dan menalisis ayat

3.2.1 Pengumpulan Ayat

Mengikut skop kajian, ayat penyata dikumpul dari buku teks BM tingkatan satu hingga tingkatan lima. Setiap ayat penyata dalam buku teks tersebut ditandakan sebelum dapat ditaip ke dalam komputer. Terdapat sebanyak 2596 ayat berjaya dikumpul mengikut skop kajian. Bagi memastikan ayat yang diskopkan sahaja yang diambil dan terbatas kepada 14 patah perkataan, setiap ayat diasingkan mengikut jumlah perkataan.

Seterusnya, ayat ini dilakar pohon sintaksis secara manual untuk mendapatkan rumus dengan berpandukan kepada buku Abrak (2005) bagi mendapatkan jenis kelas kata.

Setiap ayat terlebih dahulu ditandakan kelas kata sebelum penentuan jenis frasa dilakukan. Penentuan ini dibuat bermula dari bawah dan akhirnya membentuk lakaran pohon sintaksis secara keseluruhan.

3.2.2 Pengesahan Lakaran

Lakaran yang dibuat dinilai oleh Munsyi Dewan (MD) DBP selaku pakar BM dan pensyarah universiti yang mempunyai kemahiran dalam rumus binaan ayat BM. MD dipilih sebagai pakar BM bagi mengesahkan setiap rumus, ayat dan perkataan. MD bertanggungjawab untuk memberi ceramah dan bantuan tentang BM kepada sektor awam dan swasta (<http://munsyi.dbp.my/>). Pengesahan turut dibantu oleh Encik Nasrun bin Alias selaku pensyarah bahasa di UKM Bangi. Hasil pengesahan yang dilakukan, terdapat sebanyak 1088 ayat perlu diasingkan kerana tidak mengikut skop kajian seperti dalam Jadual 3.1.

Jadual 3.1

Jumlah Ayat yang Diasingkan

Pembahagian analisis	Jumlah ayat
Berulang	1
Ayat songsang	143
Sukar dianalisis	74
Tiada subjek	10
Klausa	312
Ayat suruhan	2
Ayat majmuk	441
Penanda wacana	104
Ayat tanya	1
Jumlah ayat	1088

Daripada jumlah 1088 ayat yang diasingkan, klausa atau ayat bukan penyata lebih banyak dikenal pasti oleh MD. Selain itu, terdapat juga beberapa ayat yang tidak dapat dianalisis struktur pembentukannya ekoran kesukaran pembentukan ayat tersebut yang termasuk dalam ayat kompleks. Terdapat 74 ayat tergolong dalam golongan ini. Menurut MD, ayat tersebut tidak boleh dianalisis berdasarkan pemahaman yang mengelirukan. Dua ayat yang lain didapati sebagai ayat suruhan. Oleh yang demikian, setelah mengetepikan sebanyak 1088 ayat berdasarkan skop kajian, selebihnya 1508 ayat dijadikan sebagai data kajian.

Oleh yang demikian, daripada 2596 ayat terkumpul, diasingkan sebanyak 1088 ayat setelah MD menentukan ayat tersebut perlu diasingkan seperti dalam Jadual 3.1. Maka data kajian ini melibatkan penggunaan 1508 ayat. Ayat ini dibahagikan kepada tiga kategori iaitu ayat untuk fasa pembangunan, fasa latihan dan fasa penilaian.

Pembahagian ayat adalah mengikut pembahagian yang dicadangkan oleh Resnik dan Lin (2013) iaitu untuk tujuan pembangunan (*development*), latihan (*training*) dan uji kaji (*testing*) prototaip. Setiap fasa mengambil kira 20% daripada data yang dikumpul untuk tujuan pembangunan, 70% untuk proses latihan dan selebihnya 10% bagi tujuan uji kaji. Oleh yang demikian, daripada 1508 jumlah ayat terkumpul kajian ini, 302 telah disediakan untuk tujuan fasa pembangunan, sebanyak 1055 untuk tujuan latihan dan selebihnya 151 adalah untuk fasa uji kaji. Pembahagian ayat ini adalah mengikut peratusan pembahagian ayat yang dimiliki oleh setiap pola ayat. Predikat yang diwakili

oleh frasa kata kerja adalah ayat yang paling banyak dikumpul iaitu 65% daripada keseluruhan ayat terkumpul seperti yang ditunjukkan dalam Jadual 3.2.

Jadual 3.2

Pembahagian Ayat

Pola Ayat	Jumlah Ayat	Peratusan	Pembangunan	Latihan	Uji kaji
Frasa nama	176	12%	36	127	18
Frasa kerja	981	65%	196	686	98
Frasa adjektif	243	16%	48	169	24
Frasa sendi nama	108	7%	22	74	11
Jumlah	1508	100%	302	1055	151

Daripada jumlah 1508 ayat terkumpul, rumus binaan ayat disenaraikan. Rumus ini dibahagikan mengikut jenis frasa. Jenis frasa diperoleh daripada hasil lakaran pohon yang dibuat pada proses sebelumnya iaitu dibahagikan mengikut pola ayat dasar iaitu Frasa nama (N"), Frasa kerja (K"), Frasa adjektif (A") dan Frasa sendi nama (KS").

3.2.3 Kumpul Atribut Perkataan

Setelah pengesahan dilakukan, semua perkataan yang terdapat dalam ayat sah dikumpul. Perkataan ditapis bagi tujuan berikut 1) tidak mempunyai perkataan yang sama supaya tiada pertindihan perkataan, dan 2) penentuan jenis kelas kata perkataan untuk dikodkan ke dalam pengaturcaraan semasa pembangunan prototaip.

Atribut bagi setiap perkataan dikumpul. Atribut ini adalah kelas kata, kata terbitan, terjemahan, imej dan ayat contoh. Atribut kata terbitan awalan dan akhiran ditentukan dengan tiga contoh mengikut skop kajian. Imej berukuran saiz passport disertakan. Terjemahan pula diambil mengikut kamus Oxford Fajar. Manakala ayat contoh pula diambil daripada buku teks BM tingkatan satu hingga tingkatan lima. Sebagai contoh bagi perkataan "rumah", kata terbitan yang disimpan adalah "perumahan", terjemahan pula adalah "house", imej pula adalah gambar rumah dan ayat contoh seperti "Rumahku di seberang sana" dan "Rumah itu kelihatan besar".

3.2.4 Reka Bentuk Pangkalan Data

Pangkalan data yang dibangunkan terdiri daripada fail rumus, fail perkataan, fail imej, fail ayat contoh, fail contoh ayat, fail ayat majmuk dan fail istilah. Fail rumus menyimpan semua rumus *X-bar* termasuk jenis kelas kata bagi setiap perkataan. Fail ini digunakan semasa padanan rumus dan kelas kata dalam membuat paparan VPS. Atribut perkataan yang terhasil daripada VPS melibatkan fail perkataan, fail imej dan fail ayat contoh. Manakala fail ayat majmuk pula digunakan semasa semakan ayat dilakukan agar hanya ayat penyata yang dianalisis.

Fail rumus adalah fail pertama yang direka bentuk. Data yang disimpan digunakan untuk reka bentuk fail perkataan. Sebagai contoh, fail perkataan akan menyimpan atribut (kelas kata, kata terbitan, terjemahan, ayat contoh) bagi setiap perkataan yang terdapat dalam fail rumus. Pautan untuk membuka fail imej juga akan disimpan dalam fail ini. Kaedah simpanan perkataan ini ditunjukkan dalam Rajah 3.4.

```

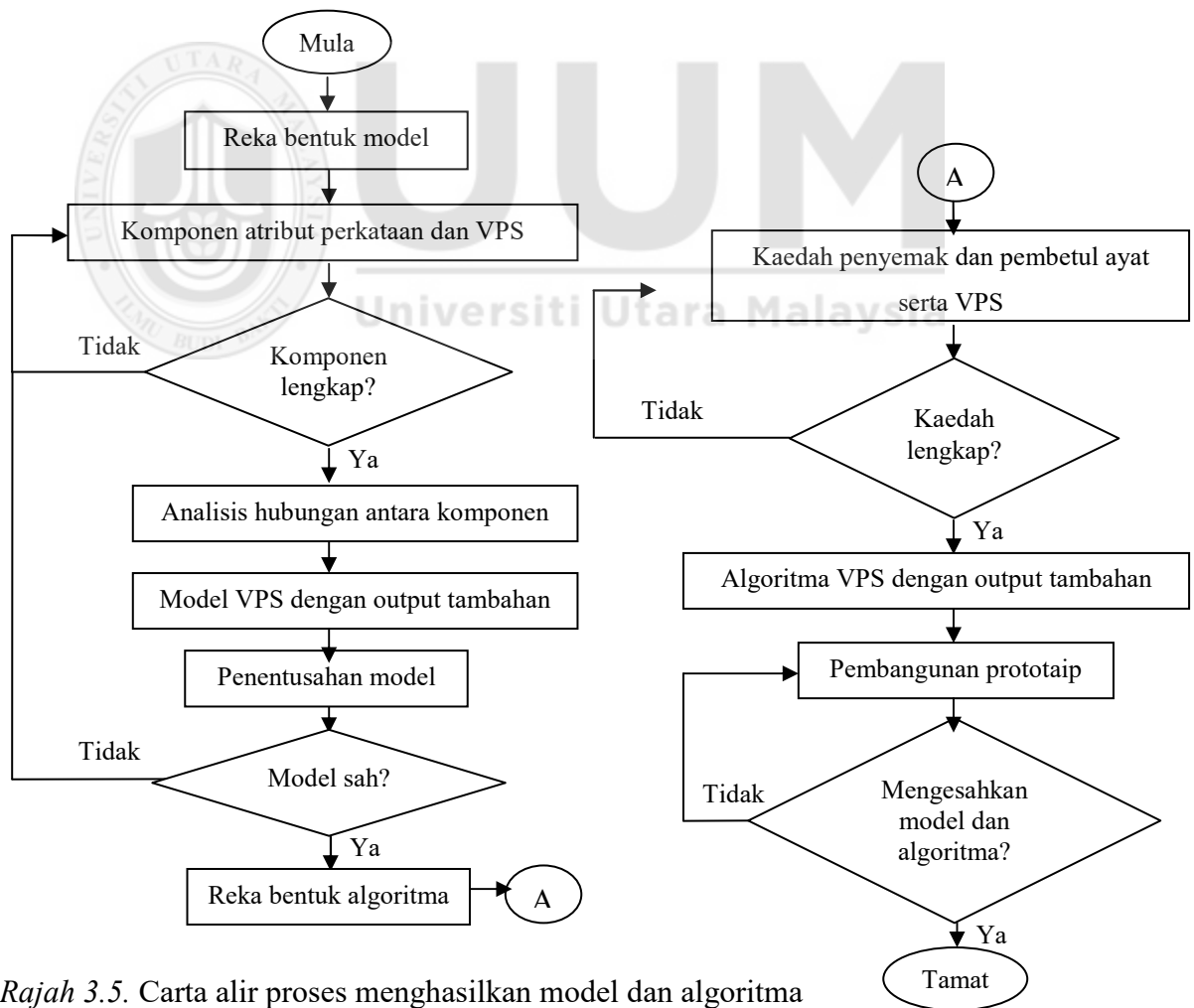
tahun | T:year | GK:'Kata Nama' | KT:'tahunan, bertahun' | I:i
masyarakat | T:society | GK:'Kata Nama' | KT:'bermasyarakat' |
melayu | T:malay | GK:'Kata Nama' | KT:'melayukan' | I:imej/me
kain | T:cloth | GK:'Kata Nama' | KT:'...' | I:imej/kain.jpg
songket | T:songket | GK:'Kata Nama' | KT:'...' | I:imej/songk
khairi | T:khairi | GK:'Kata Nama' | KT:'...' | I:imej/khairi.

```

Rajah 3.4. Perkataan disimpan dalam Fail Perkataan

3.2.5 Reka Bentuk Model dan Algoritma

Proses yang terlibat dalam fasa ini adalah untuk menghasilkan model atribut perkataan dan model VPS dengan output tambahan. Model VPS ini seterusnya digunakan untuk mereka bentuk algoritma VPS dengan output tambahan yang meliputi kaedah semakan, cadangan, dan VPS. Aktiviti yang terlibat ditunjukkan dalam Rajah 3.5.



Rajah 3.5. Carta alir proses menghasilkan model dan algoritma

Model bagi atribut perkataan disenaraikan semua atribut terlibat. Setiap atribut dirancang tentang kaedah paparan dan hubungan antara atribut. Model atribut perkataan digabung bagi mendapatkan model VPS. Gabungan antara atribut adalah merujuk kepada model SSTC dalam karya terdahulu. Model VPS ini seterusnya ditentukan oleh pakar BM dan pakar bidang pengkomputeran linguistik. Penentuan komponen dilakukan bertujuan bagi memastikan komponen terlibat mempunyai kepentingan dalam VPS ayat BM.

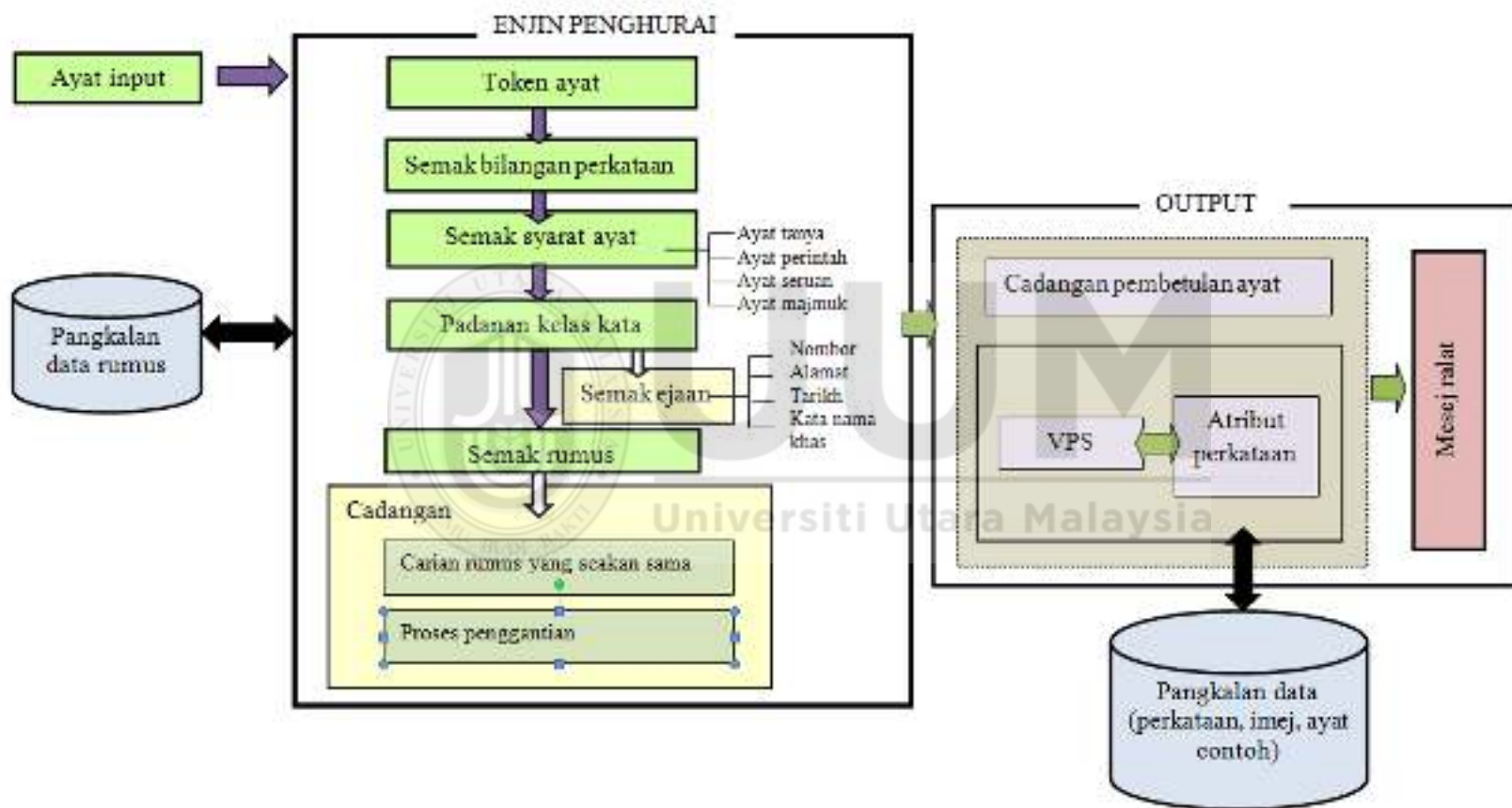
Model VPS dengan output tambahan yang dihasilkan diguna untuk mereka bentuk algoritma berkaitan. Algoritma semakan, cadangan dan VPS terlebih dahulu dihasilkan mengikut padanan dengan komponen model. Algoritma ini meliputi turutan untuk membuat semakan ayat berdasarkan padanan rumus. Status ayat yang tidak tepat, cadangan pembetulan ayat diberikan. Jenis kesalahan dikenal pasti dan ayat dengan binaan rumus yang betul dipaparkan. Seterusnya, VPS dihasilkan bagi ayat dengan semakan betul. Turutan algoritma ini dibuat dalam persamaan matematik dan carta alir yang akhirnya digabung menjadi algoritma VPS dengan output tambahan.

3.3 Fasa 3: Pembangunan Prototaip

Tujuan pembangunan prototaip adalah untuk membuat pembuktian konsep model dan algoritma. Prototaip dibangunkan menggunakan pengaturcaraan Python. Untuk menganalisis struktur ayat, pengaturcaraan Python memudahkan paparan pohon sintaksis dibuat.

Proses pertama adalah dengan membuat lakaran antara muka. Paparan ini dihubungkan dengan Fail rumus. Setelah kedua-dua bahagian dapat dihubungkan, proses seterusnya adalah atur cara pemprosesan ayat untuk membahagikan ayat kepada perkataan dan mengira jumlah perkataan. Seterusnya atur cara semakan, padanan, cadangan dan VPS diteruskan. Setiap proses mempunyai ulangan jika terdapat atur cara yang gagal semasa proses pembangunan dilakukan. Paparan output atribut perkataan diteruskan setelah VPS berjaya dipaparkan. Rajah 3.6 menunjukkan seni bina prototaip yang dilakukan.





Rajah 3.6. Seni bina prototaip VPS dengan output tambahan.

Terdapat empat jenis pangkalan data yang dibangunkan. Pangkalan data rumus digunakan untuk menyimpan semua rumus *X-bar*, frasa dan perkataan. Pangkalan data imej, perkataan dan ayat contoh pula digunakan untuk menyimpan semua komponen terlibat bagi paparan atribut perkataan. Penjelasan lanjut tentang setiap pangkalan data boleh dirujuk dalam Bab Lima.

3.4 Fasa 4: Penilaian Prototaip

Menurut Muhamad Shahbani (2012) dan Syamsul Bahrin (2011), kaedah untuk menilai hasil kajian adalah bergantung kepada kesesuaian kaedah yang dipilih dengan hasil kajian yang dilakukan. Antara kaedah yang boleh diikuti untuk menilai hasil kajian reka bentuk boleh dibuat dengan membuat pemerhatian, analitikal, eksperimen atau uji kaji, pengujian dan penilaian secara deskriptif seperti yang disenaraikan dalam Jadual 3.3.

Jadual 3.3
Kaedah dan Teknik Penilaian Reka Bentuk

Kaedah penilaian	Contoh Teknik
Pemerhatian	<ul style="list-style-type: none"> • Kajian kes - mengkaji artifak secara mendalam dalam persekitaran perniagaan • Kajian lapangan - memantau penggunaan artifak dalam pelbagai projek
Analisis	<ul style="list-style-type: none"> • Analisis statik - memeriksa struktur artifak untuk kualiti statik • Analisis seni bina- kajian mengenai kesesuaian artifak ke dalam seni bina teknikal IS • Pengoptimuman - menunjukkan ciri-ciri optimum kewujudan artifak atau menyediakan batas optimaliti pada

Sambungan Jadual 3.3

Kaedah penilaian	Contoh Teknik
	kelakuan artifak
	<ul style="list-style-type: none"> • Analisis dinamik - kajian artifak digunakan untuk kualiti dinamik
Uji kaji	<ul style="list-style-type: none"> • Eksperimen terkawal - kajian artifak dalam persekitaran terkawal bagi kualiti • Simulasi - melaksanakan artifak dengan data tiruan (<i>artificial data</i>)
Pengujian	<ul style="list-style-type: none"> • Pengujian fungsian (<i>Black box</i>) - melaksanakan antara muka artifak untuk mengenal pasti kegagalan dan kecacatan • Pengujian struktur (<i>White box</i>) - melaksanakan pengujian liputan daripada beberapa metrik (cth. laluan pelaksanaan) dalam pelaksanaan artifak
Deskriptif	<ul style="list-style-type: none"> • Hujah berdasarkan maklumat - menggunakan maklumat daripada pangkalan pengetahuan (cth. penyelidikan yang berkaitan) untuk membina satu hujah yang meyakinkan untuk utiliti artifak • Senario - membina senario terperinci sekitar artifak untuk menunjukkan utiliti

(Sumber: Muhamad Shahbani (2012) dan Syamsul Bahrin (2011), dipetik daripada Hevner *et al.* (2004))

Menurut Cleven, Gubler dan Huner (2009), kaedah yang boleh diikuti dalam kajian reka bentuk adalah seperti kajian tindakan (*action research*), kajian kes, eksperimen bidang, bukti-bukti formal, uji kaji/eksperimen terkawal, prototaip dan tinjauan. Dalam kajian ini, kaedah uji kaji atau eksperimen terkawal dipilih seperti yang dilakukan dalam kajian Muhamad Shahbani (2012) dan Syamsul Bahrin (2011). Hal ini kerana penilaian sesebuah penghurai ayat adalah satu uji kaji terkawal (*controlled experiment*) di bawah

kawalan pengkaji itu sendiri (Kakkonen, 2007 seperti yang dipetik daripada Jarvinen & Jarvinen, 2001).

Terdapat dua kaedah penilaian yang dilakukan. Kaedah penilaian Parseval dan penilaian pengguna. Penilaian Parseval dilakukan ke atas prototaip seperti kaedah yang diperkenalkan oleh Cleven *et al.* (2009) iaitu kaedah uji kaji menggunakan prototaip bagi menguji kebolehpercayaan output VPS dan cadangan pembedahan ayat. Hasil uji kaji mestilah dapat menghasilkan VPS dan memberi cadangan pembedahan ayat yang tepat. Ini bertujuan untuk mengetahui kegagalan dan kecacatan hasil yang dipaparkan yang menunjukkan model dan algoritma perlu ditambah baik. Seterusnya, penilaian pengguna yang juga termasuk dalam uji kaji terkawal dilakukan dengan pengguna akhir untuk mendapatkan maklum balas kepuasan subjektif dan kognitif. Kaedah penilaian ini dijelaskan dalam bahagian seterusnya.

3.4.1 Kaedah Penilaian Parseval

Teknik penilaian formatif atau summatif boleh diikuti untuk menilai output yang dihasilkan oleh prototaip (Resnik & Lin, 2013). Teknik formatif dibuat ke atas prototaip semasa fasa latihan untuk mendapatkan rumus yang tepat. Manakala teknik summatif dilakukan semasa fasa uji kaji. Teknik penilaian ini adalah menurut kaedah intrinsik iaitu kaedah menilai fungsi prototaip dari segi output yang dihasilkan.

Metrik penilaian dalam pemprosesan bahasa tabii atau sesebuah penghurai ayat dikenali sebagai Parseval yang diukur menggunakan metrik penilaian *recall*, *precision* dan *f-score* atau juga dikenali sebagai *f-measure*. Metrik penilaian ini digunakan untuk

mengukur bilangan output penghurai bagi setiap ayat input mengikut ukuran penghurai "*gold-reference* atau *gold-standard*" atau juga dikenali sebagai "*reference output*" iaitu menganggarkan kewujudan penghurai ayat yang dilabelkan oleh manusia dan dibandingkan dengan output yang diperoleh melalui sistem. Output ini boleh dibandingkan dari segi hasil konstituen, kebergantungan atau ayat. Selain daripada itu, dalam ukuran metrik Parseval, penilaian juga boleh diukur dengan membandingkan output simbol bracket (Resnik & Nik, 2013; Noor Hafhizah, 2011; Powers, 2011; Jurafsky & Martim, 2009; Kikas & Treumuth, 2007; Kakkonen, 2007; Hirschman & Mani, 2004; Manning & Schutze, 2000).

Kakkonen (2007) menjelaskan bahawa penilaian sesebuah penghurai ayat boleh diukur dari segi ketepatan output (*correctness or preciseness of the output*), kecekapan (*efficiency*) dan kebolehgunaan (*usability*). Ketepatan output dinilai dengan memeriksa hasil ketepatan output yang dihasilkan oleh sistem menggunakan kaedah "*gold standard*" iaitu dengan membuat perbandingan dengan jawapan yang telah disediakan. Kecekapan pula dinilai dari segi masa dan ruang, manakala kebolehgunaan pula dinilai oleh pengguna sistem. Perspektif yang biasa digunakan dalam penilaian penghurai ayat adalah ketepatan (*precision*) output yang dihasilkan.

Menurut Kovar (2014) dan Resnik dan Lin (2013), paradikma penilaian dalam NLP biasanya melibatkan ukuran *recall* dan *precision*. Namun, pengkaji boleh menggunakan satu ukuran metrik penilaian sahaja untuk memberi rumusan ke atas hasil output yang dilakukan. Oleh yang demikian, metrik penilaian *recall* dan *precision* digunakan dalam

kajian ini. Pengkaji NLP memberi maksud *recall* sebagai hasil output penghurai yang berjaya dikeluarkan oleh sistem, manakala maksud *precision* pula adalah hasil ketepatan penghurai yang dapat diperoleh daripada hasil *recall*. Contoh ringkasan maksud ini ditunjukkan dalam Jadual 3.4.

Jadual 3.4

Metrik Penilaian Penghurai Ayat

Metrik penilaian	Resnik dan Lin (2013)	Noor Hafhizah (2011) dipetik daripada Carroll et al. (1998)	Kikas dan Treumuth (2007)
<i>Recall</i>	Jumlah output dihasilkan	$Recall = \frac{Parsed\ sentence\ (Jumlah\ output)}{Intended\ parsed\ sentence}$ (Output yang dikira)	$\frac{Output\ constituents}{Constituents\ in\ parser\ output}$
<i>Precision</i>	Mengukur tahap ketepatan output yang dihasilkan	$Precision = \frac{Output\ tepat\ (correct\ parsed\ sentence)}{Semua\ ayat\ uji\ kaji\ (all\ parsed\ sentence)}$	$\frac{Correct\ constituents}{Constituents\ in\ gold\ standard}$
<i>F-Score / F-measure</i>	Skor yang menggabungkan <i>recall</i> dan <i>precision</i> menjadi angka tunggal	$(precision + recall) / 2$	Imbangan antara <i>recall</i> dan <i>precision</i>

Menurut Bastings dan Sima'an (2014), penilaian output yang diukur berdasarkan output pohon sintaksis, persamaan di antara output pohon sintaksis dengan output pohon yang dilabelkan oleh manusia boleh dilakukan. Perbezaan bentuk di antara kedua-dua pohon sintaksis perlu dilakukan. Persamaan dan perbezaan label kelas kata juga memberi kesan ke atas ketepatan pohon sintaksis yang dikeluarkan. Ketepatan output ini dikira berdasarkan kepada bentuk pohon sintaksis, perwakilan jenis frasa dan kelas kata yang dikeluarkan dibandingkan dengan jawapan dalam *gold standard*. Oleh yang demikian, output yang dihasilkan dalam kajian ini dibandingkan dengan output lakaran yang telah

disahkan oleh Munsyi Dewan sebelumnya. Bagi mendapatkan keputusan uji kaji yang lebih tepat, setiap output dinilai ketepatannya oleh Munsyi Dewan Puan Siti Salmah binti Sulaiman dan perbincangan turut dilakukan bersama Encik Nasrun bin Alias sebagai Pensyarah Kanan di Universiti Kebangsaan Malaysia dalam bidang Bahasa Melayu.

3.4.2 Kaedah Penilaian Pengguna

Penilaian artifak (model atau prototaip) boleh dibuat secara eksperimen terkawal melalui penilaian pengguna (Muhamad Shahbani, 2012 dipetik daripada Hevner, 2004; Rubin & Chisnell, 2008). Penilaian ini adalah untuk mengukur tahap kebolegunaan (*usability*). Kebolegunaan dikaitkan dengan kebergunaan (*usefulness*) fungsi sistem yang boleh dimanfaatkan oleh pengguna (Nielsen, 1993). Tiga jenis teknik penilaian kebolegunaan iaitu penilaian analisis (*analytic evaluation*), penilaian oleh pakar (*evaluation by experts*), dan penilaian oleh pengguna (*evaluation by users*). Penilaian oleh pengguna adalah untuk mengukur tahap kepuasan subjektif (Tullis & Albert, 2013; Laventhal & Barnes, 2008). Penilaian ini boleh diukur menggunakan soal selidik (Nielsen, 1993). Kajian ini mengadaptasi soal selidik *Usefulness, Satisfaction, and Ease of use* (USE) iaitu soal selidik untuk menilai tahap kepuasan subjektif pengguna (Lund, 2001) yang dirujuk daripada <http://garyperلمان.com/que>

Persampelan

Berdasarkan domain kajian ini, sampel yang terlibat adalah 30 orang pelajar tingkatan satu hingga tingkatan lima dari Sekolah Menengah Kebangsaan Changlun, Kedah. Sekolah ini dipilih kerana berdekatan dengan tempat kajian. Saiz sampel ditentukan

menurut Nielsen (1993) yang menyatakan bahawa penggunaan soal selidik sebaiknya melibatkan 30 atau lebih responden. Walaupun Tullis dan Albert (2013) dan Nielsen (2000) menyatakan bahawa lima orang responden adalah memadai. Oleh kerana responden yang terlibat adalah pelajar tingkatan satu hingga tingkatan lima, maka pembahagian pemilihan sampel adalah enam orang untuk setiap tingkatan.

Prosedur persampelan rawak mudah (*simple random sampling*) dipilih dalam kajian ini. Hal ini kerana, mengikut domain kajian, sampel yang terlibat telah diketahui, maka kaedah persampelan kebarangkalian diikuti (Scheaffer, Mendenhall III & Ott, 2006). Dalam kaedah ini, pemilihan sampel boleh dilakukan secara rawak jika sampel tidak akan terlibat melebihi sekali dan kebarangkalian pemilihan juga sama rata iaitu berkemungkinan terpilih (Barnett, 2002). Kaedah soal selidik dalam pengumpulan data juga adalah kaedah yang digunakan dalam prosedur persampelan rawak mudah dan kebarangkalian. Oleh yang demikian, responden dipilih adalah secara rawak mengikut kesediaan pada masa soal selidik dilakukan. Persetujuan daripada responden juga diperolehi (Lampiran B).

Oleh kerana penilaian ini melibatkan pelajar sekolah menengah, maka strategi pembelajaran perlu diambil kira. Strategi pembelajaran ini melibatkan strategi kognitif (Azizah, 2012, dipetik daripada Somuncuoglu & Yildirim, 1999). Kajian ini juga melibatkan penggunaan teori beban kognitif, maka penilaian kognitif perlu dilakukan. Instrumen penilaian strategi kognitif dinilai dengan mengguna soal selidik yang diadaptasi daripada Somuncuoglu dan Yildirim (1999).

Instrumen penilaian

Soal selidik yang digunakan dibahagikan kepada tiga bahagian. Bahagian A adalah untuk mendapatkan maklumat responden dari segi umur, jantina dan pengetahuan tentang pohon sintaksis. Maklumat yang diperoleh mendapati 29 orang responden mempunyai pengetahuan tentang pohon sintaksis. Sebanyak 20 orang daripada responden adalah pelajar lelaki. Bahagian B merangkumi 30 soalan yang diadaptasi daripada soal selidik USE. Soalan ini dibahagikan kepada empat dimensi iaitu kebergunaan (*usefulness*), kemudahan penggunaan (*ease of use*), kemudahan pembelajaran (*ease of learning*) dan kepuasan subjektif (*subjective satisfaction*). Bahagian C pula terdiri daripada 10 soalan berkenaan penilaian kognitif. Soalan daripada soal selidik USE menggunakan skala perbezaan semantik 1 (sangat tidak setuju) hingga skala 7 (sangat setuju). Manakala soalan penilaian kognitif menggunakan skala likert 1 (sangat tidak setuju) hingga skala 5 (sangat setuju). Oleh kerana soal selidik yang digunakan diadaptasi daripada penulis berbeza, maka skala yang digunakan tidak diubah walaupun skala 1-7 dan skala 1-5 digunakan. Hal ini juga kerana tujuan penilaian kepuasan dan kognitif juga berbeza.

Instrumen yang digunakan walaupun diadaptasi daripada pengkaji terdahulu, namun instrumen tersebut dialih bahasa daripada BI ke BM. Ayat yang digunakan dalam setiap soalan juga disesuaikan dengan pemahaman pelajar sekolah menengah namun maksud asal tetap dikekalkan. Oleh itu, instrumen ini dinilai tahap kebolehpercayaan dan kesahan dengan nilai $\alpha = 0.961$ selepas diuji dengan data terkumpul menggunakan

perisian *Statistical Package for Social Science* (SPSS). Maklumat ini ditunjukkan dalam Jadual 3.5.

Jadual 3.5

Hasil Kebolehppercayaan Instrumen Kajian

Alpha Cronbach	Alpha Cronbach berdasarkan item standard	Jumlah (N) item
.961	.963	30

Pengumpulan data

Persetujuan daripada pengetua sekolah terlebih dahulu diperolehi (Lampiran C) dengan mengemukakan surat kebenaran pengumpulan data oleh pihak universiti (Lampiran D). Soal selidik dilakukan dengan bantuan pihak sekolah untuk mengumpul responden di kawasan yang disediakan. Berdasarkan kaedah yang dilakukan Syamsul Bahrin (2011), responden terlebih dahulu diberi penerangan tentang soal selidik yang dilakukan termasuklah mendapatkan persetujuan mereka. Setelah itu, demo penggunaan prototaip ditunjukkan oleh pengkaji. Seterusnya, responden diberi peluang untuk menggunakan prototaip berpandukan kepada tugas yang telah disediakan. Tugas yang disediakan meminta pelajar memasukkan ayat mudah, membuat paparan VPS, membuat pautan dan mencuba ayat sederhana dan kompleks. Sebelum responden mengisi soal selidik, penerangan sekali lagi diberi bagi memastikan responden memahami skala yang digunakan. Seterusnya responden diberi kebebasan tanpa pengawasan pengkaji untuk mengisi soal selidik.

Analisis data

Data yang diperolehi dianalisis menggunakan perisian SPSS versi 22.0. Metrik pengiraan statistik deskriptif digunakan untuk membuat rumusan hasil yang diperolehi. Penentuan ini adalah mengikut Tullis dan Albert (2008) yang menyatakan bahawa penilaian menggunakan skala perlu diukur menggunakan statistik deskriptif. Skor min adalah antara skor yang perlu dinyatakan sebagai hasil penilaian.

3.5 Rumusan Bab Tiga

Metodologi PR dijadikan panduan dalam menjalankan kajian ini yang dibahagikan kepada lima fasa iaitu fasa 1) pengetahuan pernyataan masalah, 2) mengkategorikan dan menganalisis ayat untuk mendapatkan rumus binaan ayat, reka bentuk model dan algoritma, 3) pembangunan prototaip, 4) penilaian dan 5) rumusan. Fasa pertama membawa kepada pengukuhan pernyataan masalah, objektif dan skop kajian. Fasa kedua membawa kepada pengumpulan ayat, rumus dan reka bentuk model serta algoritma. Fasa ketiga diteruskan untuk membina prototaip kajian bagi melakukan pembuktian konsep model dan algoritma. Seterusnya, kaedah untuk menjalankan fasa penilaian dirumuskan dengan menetapkan kaedah berbentuk eksperimen prototaip yang dibahagikan kepada kaedah penilaian Parseval dan kaedah penilaian pengguna dalam Fasa keempat. Rumusan kajian dijelaskan dalam Fasa kelima termasuk perbincangan mengenai pencapaian objektif kajian.

Bab ini juga membincangkan secara ringkas tentang kaedah mendapatkan rumus binaan ayat. Rumus ini diperolehi hasil daripada lakaran pohon sintaksis yang dibuat secara

manual. Lakaran tersebut disahkan oleh pakar BM dan pakar bidang pengkomputeran linguistik. Lakaran yang disahkan diambil dan dikumpul atribut perkataan terlibat. Atribut ini disimpan dalam pangkalan data mengikut rumus, perkataan, ayat contoh, imej dan ayat majmuk.

Reka bentuk model dan algoritma juga dibincangkan secara ringkas. Pembangunan model dimulai dengan menyenaraikan semua komponen bagi atribut perkataan dan VPS. Seterusnya komponen ini dilakar untuk mendapatkan model VPS berserta output tambahan. Setelah model ditentukan, reka bentuk algoritma dimulai dengan membuat semakan serta cadangan ayat sebelum algoritma VPS dengan output tambahan dihasilkan.

Selain itu, kaedah pembangunan prototaip dihuraikan yang meliputi lakaran antara muka dan pengkodan atur cara bagi proses semakan, cadangan dan VPS. Prototaip ini digunakan untuk menjalankan fasa penilaian. Tujuan penilaian ini adalah untuk mengesahkan model dan algoritma yang direka bentuk. Penjelasan secara terperinci reka bentuk model dan algoritma ini dibincangkan dalam bahagian seterusnya.

BAB EMPAT

MODEL DAN ALGORITMA

4.0 Pengenalan

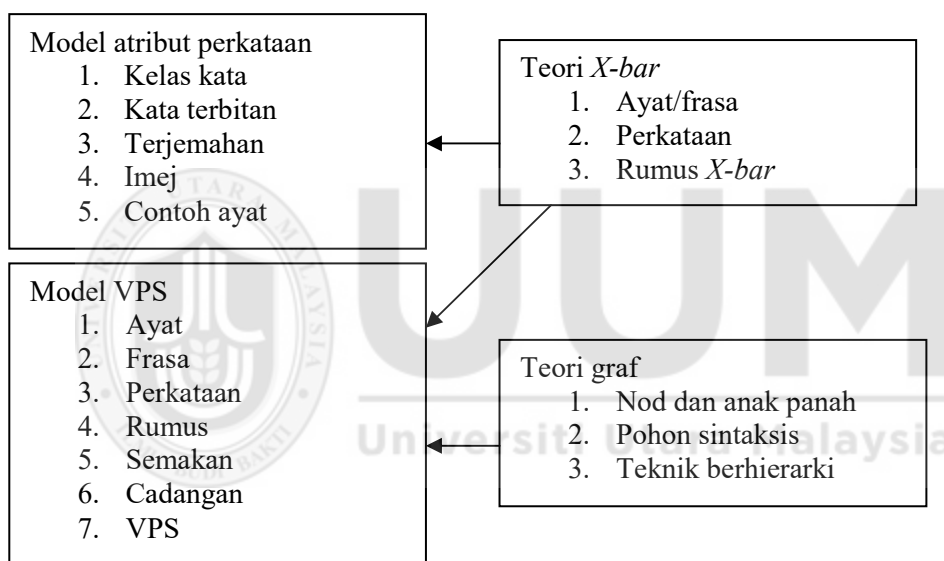
Model mengandungi komponen, perkaitan antara komponen dan persembahan (Jurafsky & Martim, 2009). Algoritma pula menerangkan prosedur perkomputeran tertentu untuk mencapai perkaitan antara input dan output (Cormen *et al.*, 2001). Penjelasan makna model dan algoritma ini dijadikan panduan untuk mereka bentuk model dan algoritma VPS serta output tambahan.

Pembangunan model dimulai dengan menyenaraikan semua komponen bagi atribut perkataan. Seterusnya komponen ini dilakar untuk mendapatkan model VPS. Setelah model ditentusahkan, reka bentuk algoritma dimulai dengan membuat semakan serta cadangan ayat sebelum algoritma VPS dengan output tambahan dihasilkan.

Bab ini memberi penjelasan untuk mencapai objektif kedua iaitu "Membina model dan algoritma VPS dengan output tambahan serta prototaip kajian sebagai alat pembuktian konsep". Penjelasan mengenai pembangunan prototaip dijelaskan dalam Bab Lima. Secara ringkas, Bab ini membuat huraian mengenai penghasilan 1) model atribut perkataan, 2) model VPS berserta output tambahan dan penentusahan model, dan 3) algoritma VPS dengan output tambahan. Model dan algoritma VPS dengan output tambahan yang dihasilkan adalah objektif utama kajian ini. Setiap proses pembangunan ke atas ketiga-tiga keperluan ini dijelaskan dalam bahagian seterusnya.

4.1 Pembangunan Model

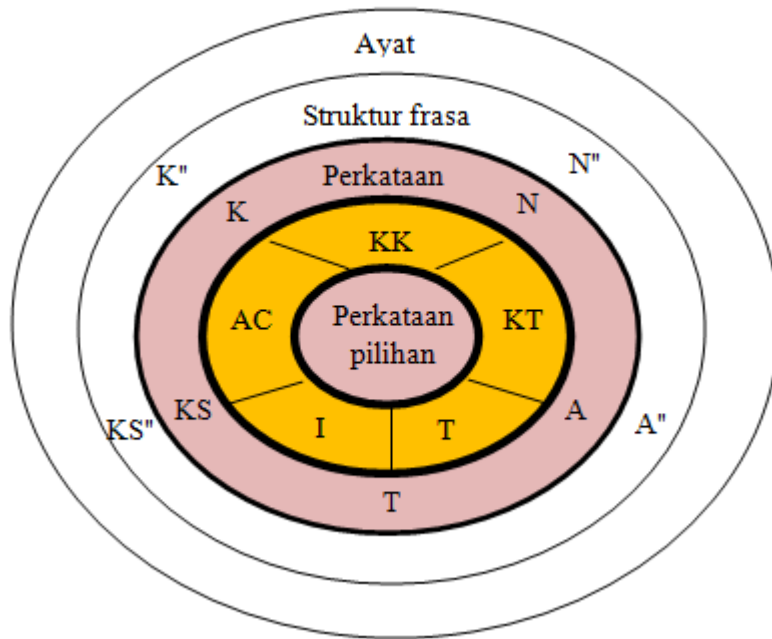
Terdapat dua jenis model yang dibangunkan iaitu model atribut perkataan dan model VPS dengan output tambahan. Komponen yang terdapat dalam setiap model adalah berdasarkan kepada teori terlibat. Komponen ini adalah ayat/frasa, perkataan dan rumus sebagaimana yang terdapat dalam teori *X-bar*. Komponen ini dipaparkan dalam VPS berbentuk VPS secara hierarki yang mempunyai nod dan anak panah seperti yang terdapat dalam teori graf. Perkaitan ini ditunjukkan dalam Rajah 4.1.



Rajah 4.1. Komponen model dan teori

4.1.1 Model Atribut Perkataan

Model atribut perkataan (Rajah 4.2) mengandungi lima komponen iaitu kelas kata (KK), kata terbitan (KT), terjemahan (T), imej (I) dan senarai ayat contoh (AC). Analisis yang dilakukan ke atas KK, KT, dan T mengikut teori *X-bar* adalah di bawah skop komponen perkataan. Manakala ayat contoh pula adalah komponen ayat dan rumus.



Rajah 4.2. Model atribut perkataan

Rajah 4.2 menunjukkan satu perkataan boleh menghasilkan KK, KT, T, I, dan AC dalam model berbentuk bulatan bagi menunjukkan perkaitan antara setiap komponen. Perkaitan ini menunjukkan bahawa setiap atribut adalah mencerminkan komponen bagi perkataan yang sama. Perkataan pilihan terbit daripada struktur frasa yang terdapat dalam VPS yang berakar daripada sebuah ayat.

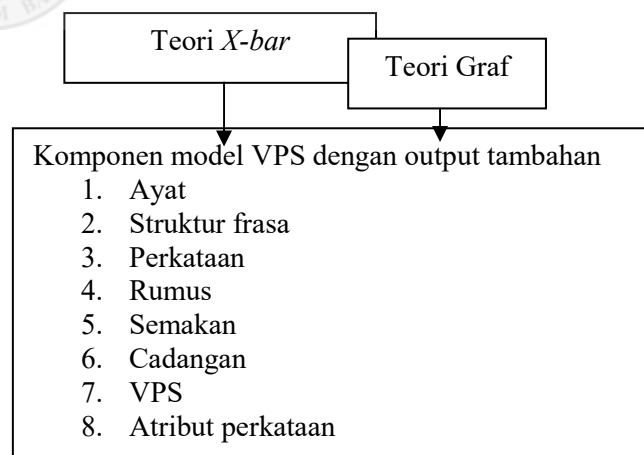
4.1.2 Model VPS dengan Output Tambahan

Dalam VPS yang dihasilkan dalam kajian ini, semua jenis kelas kata dikaji. Dalam BM jenis kelas kata dibahagikan kepada kata nama (N), kata kerja (K), kata adjektif (A), kata sendi nama (KS) dan jenis kata yang lain dipanggil sebagai kata tugas (T). Sama juga dengan jenis frasa yang terlibat iaitu frasa nama (N''), frasa kerja (K''), frasa adjektif

(A") dan frasa sendi nama (KS"). Gabungan jenis kelas kata dan jenis frasa ini akan menghasilkan struktur pembentukan ayat yang lengkap.

Gabungan jenis kelas kata dan frasa memerlukan komponen lain untuk menghurai struktur ayat. Komponen tersebut perlu mengandungi perkataan/frasa/ayat, leksikon/pangkalan data, penghurai ayat, dan rumus tatabahasa/leksikal. Komponen ini ditambah baik dengan beberapa penambahan komponen berdasarkan skop kajian. Oleh yang demikian, komponen model VPS dengan output tambahan yang dikemukakan adalah 1) **ayat**, 2) **struktur frasa**, 3) **perkataan**, 4) **rumus *X-bar***, 5) **semakan ayat**, 6) **cadangan ayat**, 7) **VPS**, dan 8) **atribut perkataan**.

Perkaitan antara keperluan komponen model VPS serta output tambahan dengan komponen teori ditunjukkan melalui Rajah 4.3.

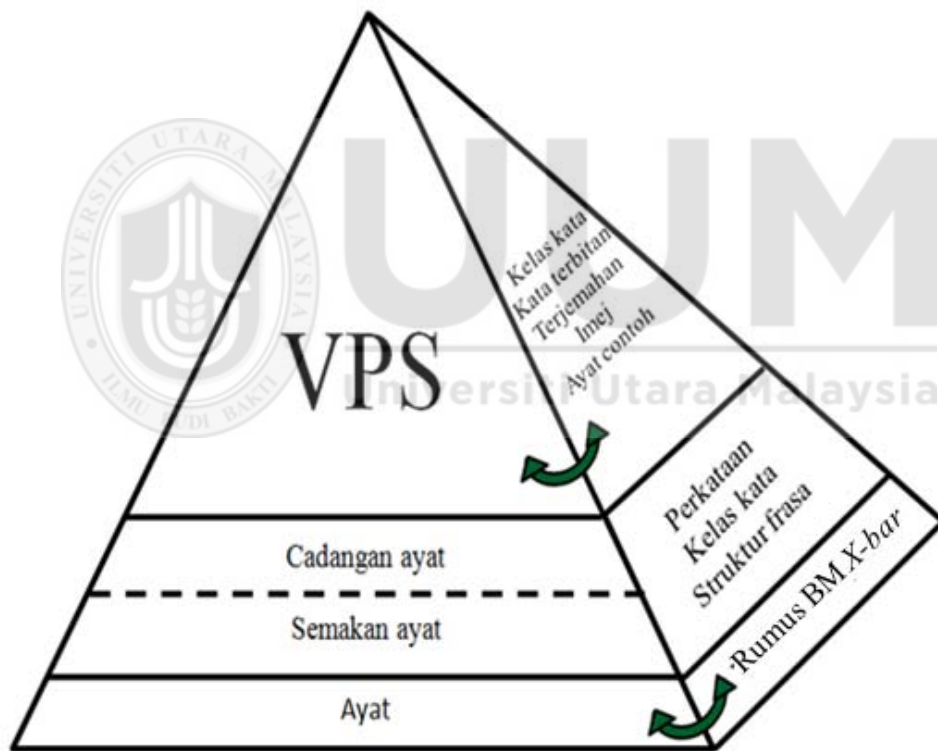


Rajah 4.3. Perkaitan komponen model VPS dengan output tambahan dan teori

Komponen perkataan, frasa, ayat dan Rumus *X-bar* digabung dalam komponen semakan dan cadangan ayat. Hal ini kerana semakan dan cadangan melibatkan analisis ke atas

perkataan hinggalah struktur pembentukan mengikut rumus. Oleh itu, model VPS melibatkan komponen semakan ayat, cadangan ayat, VPS, dan atribut perkataan.

Semakan dan cadangan pembedahan ayat terlebih dahulu dilakukan sebelum dapat menghasilkan VPS. Dalam VPS, visualisasi memaparkan jenis kelas kata, jenis frasa dan atribut perkataan ayat input. Senarai ayat contoh dalam atribut perkataan mempunyai pautan untuk membuat VPS yang baharu di mana proses ulangan dalam pembentukan VPS dilakukan seperti dalam Rajah 4.4.



Rajah 4.4. Model VPS dengan output tambahan (model piramid)

Rajah 4.4 menunjukkan model VPS berbentuk piramid. Konsep model berbentuk piramid dipilih kerana konsep ini jelas untuk menunjukkan proses dilakukan mengikut turutan yang bermula daripada bawah. Turutan ini menunjukkan proses semakan ayat

melakukan padanan ke atas perkataan, kelas kata dan struktur frasa. Cadangan pembetulan pula akan membuat proses penggantian ke atas atribut tersebut. Antara kedua-dua proses ini, garisan bersambung digunakan sebagai maksud proses pilihan (*options*) yang ada. Padanan komponen dan fungsi yang terlibat ditunjukkan dalam Jadual 4.1.

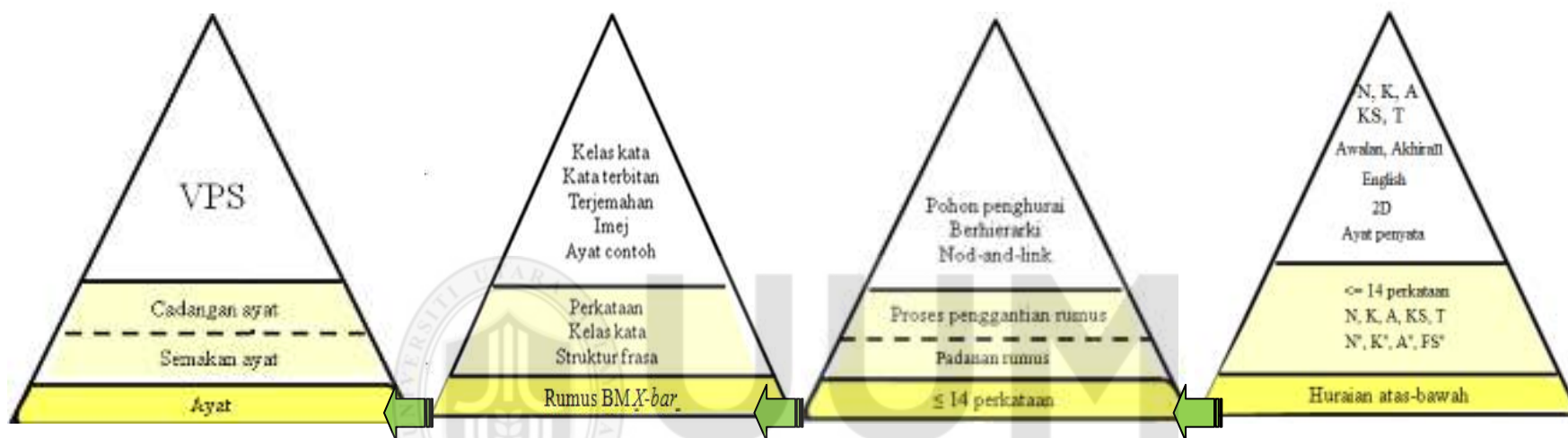
Jadual 4.1

Komponen Model VPS dengan Output Tambahan

Komponen	Fungsi
1. Semakan ayat	Semakan ayat melakukan padanan kelas kata dan semakan rumus.
2. Cadangan ayat	Cadangan ayat diberikan bagi ayat yang tidak tepat mengikut rumus. Cadangan ini berdasarkan penggantian perkataan ayat input mengikut rumus yang disimpan.
3. VPS	Pohon sintaksis dipaparkan dalam bentuk hierarki. VPS dihasilkan berdasarkan kepada ayat input manakala VPS dari ayat contoh adalah daripada prototaip. <ul style="list-style-type: none"> A. VPS ayat input <ul style="list-style-type: none"> Hanya dipaparkan jika ayat yang dimasukkan adalah betul mengikut rumus. B. VPS ayat contoh <ul style="list-style-type: none"> Senarai ayat contoh diberikan dalam paparan atribut perkataan. Paparan ini berdasarkan pemilihan perkataan dalam VPS. Setiap ayat disertakan dengan pautan untuk membuat VPS yang baharu.
4. Atribut perkataan	Setiap atribut disenaraikan mengikut perkataan yang dipilih pada pautan VPS ayat input. Ayat contoh dalam konteks yang berbeza juga disenaraikan berserta pautan untuk membuat VPS baharu.

Disebalik dua sisi model piramid (depan dan kanan), terdapat ciri-ciri lain disebalik komponen utama. Setiap peringkat pemprosesan yang bermula daripada bawah, diperolehi hasil daripada komponen utama. Contohnya, komponen VPS mempunyai ciri-ciri pohon penghurai berbentuk hierarki dan mempunyai nod dan anak panah. Setiap perkaitan antara komponen utama diperincikan seperti Rajah 4.5.





Rajah 4.5. Empat sisi model piramid

Rajah 4.5 menunjukkan model piramid mempunyai tiga peringkat pemprosesan yang bermula daripada bawah iaitu ayat, semakan dengan cadangan dan VPS. Perkaitan antara setiap peringkat pemprosesan dan sisi model piramid dijelaskan melalui Jadual 4.2.

Jadual 4.2

Peringkat Pemprosesan Model Piramid

Peringkat pemprosesan	Huraian
Ayat	Ayat yang diterima adalah ayat nyata bagi perkataan kurang atau sama 14 perkataan. Syarat penerimaan ayat dan setiap perkataan dalam ayat disimpan dalam bentuk rumus <i>X-bar</i> .
Semakan dan cadangan	<p>Ayat yang diterima disemak rumus yang terlibat. Semakan dibuat dengan membuat padanan setiap perkataan, kelas kata dan struktur frasa.</p> <p>Perkataan yang disemak mestilah dalam lingkungan kurang atau sama 14 perkataan. Kelas kata yang disemak meliputi kata nama, kata kerja, kata adjektif dan kata sendi nama. Manakala struktur frasa yang terlibat adalah frasa nama, frasa kerja, frasa adjektif dan frasa sendi nama.</p> <p>Padanan yang tidak dapat dibuat diteruskan untuk membuat proses penggantian rumus bagi ketiga-tiga kriteria iaitu perkataan, kelas kata dan struktur frasa. Kriteria yang diproses juga mestilah bagi ayat kurang atau sama 14 perkataan.</p>
VPS	VPS yang dihasilkan setelah padanan rumus, adalah pohon penghurai berbentuk hierarki yang mempunyai nod dan anak panah. Setiap nod mempunyai pautan untuk mengetahui atribut perkataan yang terdiri daripada kelas kata, kata terbitan, terjemahan, imej dan senarai ayat contoh.

4.1.3 Penentusahan Model VPS dengan Output Tambahan

Penentusahan merupakan proses penilaian komponen untuk memastikan komponen yang terlibat memenuhi spesifikasi output (Pathiah, 2012). Bermatlamat untuk memastikan output yang dihasilkan akan menyediakan jawapan dan penyelesaian seperti yang diperlukan (Barr & Klavan, 2001). Penilaian pakar (*expert review*) adalah salah satu cara penentusahan model (Pathiah, 2012) seperti yang dilakukan dalam penentusahan model kajian Syamsul Bahrin (2011) dan Muhamad Shahbani (2012).

Sebanyak tiga hingga lima orang pakar bidang adalah memadai untuk memberi maklum balas (Syamsul Bahrin, 2011 dipetik daripada Schneiderman, 1998). Pemilihan pakar bidang ini dibuat berdasarkan komponen yang terdapat dalam model VPS. Ini kerana penentusahan yang dilakukan adalah melibatkan penentusahan komponen (Pathiah, 2012). Komponen yang terlibat adalah komponen struktur pembentukan ayat dalam pemprosesan BM. Justeru itu, tiga orang Munsyi Dewan sebagai pakar BM dan dua orang pakar dalam bidang perkomputeran linguistik dipilih untuk membuat penentusahan model VPS.

Berdasarkan kaedah penentusahan kajian Muhamad Shahbani (2012) dan Syamsul Bahrin (2011), maka empat kaedah penentusahan diambil kira. Kaedah ini adalah 1) penerangan mengenai tujuan penentusahan, 2) pakar menjalankan penilaian, 3) analisis dapatan, dan 4) membuat penambahbaikan model.

Atribut penilaian yang digunakan dalam kajian Muhamad Shahbani (2012) dirujuk dan diubah berdasarkan kesesuaian kajian. Kriteria yang digunakan adalah bersesuaian dengan kriteria yang perlu dinilai dalam pemrosesan bahasa iaitu antaranya penilaian tentang kandungan, struktur, dan tatabahasa (Barr, 2003). Kriteria penilaian yang merangkumi struktur dan kandungan dikategorikan sebagai kriteria secara umum. Manakala kriteria penilaian tatabahasa pula merangkumi komponen semakan ayat, cadangan ayat, VPS dan atribut perkataan. Instrumen penilaian ditunjukkan dalam Lampiran E.

Instrumen penilaian berbentuk soal selidik yang digunakan mengandungi 10 soalan menggunakan skala likert 1 (sangat tidak setuju) hingga skala 5 (sangat setuju). Soalan 1 hingga soalan 6 adalah kriteria penilaian umum, manakala kriteria penilaian ketujuh hingga sepuluh adalah kriteria penilaian untuk komponen model. Hasil skor yang diperolehi dikira nilai peratusan berdasarkan jumlah keseluruhan skor yang diperolehi untuk setiap soalan. Hasil yang diperolehi ditunjukkan dalam Jadual 4.3.

Jadual 4.3

Hasil Penentusahan Model

Soalan (S)/ Penilai (P)	P1	P2	P3	P4	P5	Jumlah	Peratus
S1	4	4	4	4	4	20	80%
S2	4	4	4	4	4	20	80%
S3	4	4	4	4	4	20	80%
S4	4	4	4	4	4	20	80%
S5	4	4	4	4	4	20	80%
S6	4	4	5	4	4	21	84%
Purata peratusan							80.6%

Para penilai bersetuju dengan kriteria penilaian umum yang diberikan dengan purata penilaian 80.6%. Kriteria tersebut adalah S1) model ini senang difahami, S2) turutan dan proses yang terlibat adalah jelas, S3) model ini bersesuaian dengan VPS ayat BM, S4) model ini memberi bantuan pemahaman dan pembelajaran ayat BM dengan paparan struktur frasa, kelas kata dan perkataan, S5) model ini memberi sokongan mendatang terhadap keperluan aplikasi NLP, dan S6) komponen model adalah bersesuaian untuk membuat VPS ayat BM. Selain daripada itu, salah seorang penilai memberi pendapat bahawa beliau sangat bersetuju dengan model VPS bagi kriteria penilaian umum.

Kriteria penilaian ketujuh adalah S7) komponen semakan ayat diperlukan dalam pembelajaran ayat secara VPS, S8) komponen cadangan ayat diperlukan dalam VPS untuk memberi cadangan pembedaan ayat mengikut rumus, S9) komponen VPS boleh

membantu pemahaman struktur ayat, kelas kata dan perkataan, dan S10) set atribut perkataan (kelas kata, kata terbitan, terjemahan, imej, ayat contoh) diperlukan dalam VPS untuk memberi pemahaman yang lebih mendalam tentang struktur ayat dan perkataan serta boleh difahami dalam konteks ayat yang lain.

Komponen semakan ayat (S7) dan VPS (S9) dinilai sebagai komponen yang sangat diperlukan dalam pembelajaran struktur ayat BM berbentuk VPS. Kedua-dua komponen mendapat penilaian 88%. Dua orang penilai menyatakan bahawa mereka sangat bersetuju dengan kedua-dua komponen ini. Manakala komponen cadangan ayat (S8) dan atribut perkataan (S10) dipersetujui untuk digunakan dengan beberapa penambahbaikan. Kedua-dua komponen mendapat penilaian sebanyak 84%. Ringkasan hasil penilaian komponen dihuraikan dalam Jadual 4.4.

Jadual 4.4

Hasil Penilaian Komponen

Komponen	Hasil penilaian
Semakan ayat	<ul style="list-style-type: none"> • Komponen semakan ayat dinilai sebagai komponen yang sangat diperlukan untuk pembelajaran ayat secara VPS. • Dua orang penilai telah memberi penilaian sebagai sangat bersetuju untuk memasukkan semakan dalam pembelajaran ayat secara VPS.
Cadangan ayat	<ul style="list-style-type: none"> • Komponen cadangan pembetulan ayat dinilai sebagai komponen yang diperlukan. Para penilai menyatakan persetujuan mereka untuk memasukkan cadangan pembetulan

Sambungan Jadual 4.4

Komponen	Hasil penilaian
	ayat mengikut rumus binaan ayat BM.
VPS	<ul style="list-style-type: none">• Para penilai bersetuju bahawa komponen VPS boleh membantu pemahaman struktur ayat, kelas kata dan perkataan. Dua orang penilai sangat bersetuju dengan komponen ini.
Atribut perkataan	<ul style="list-style-type: none">• Kelima-lima atribut perkataan juga dipersetujui oleh para penilai.

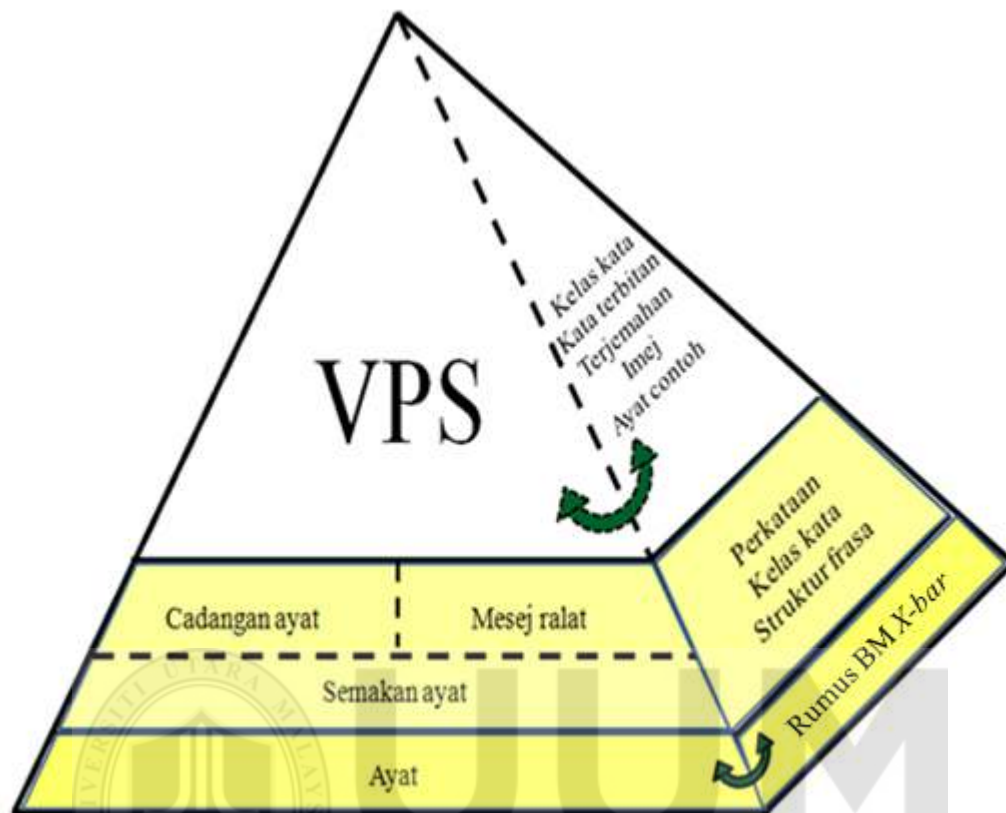
Secara keseluruhan, Prof. Madya Dr. Wan Amizah binti Wan Mahmud iaitu salah seorang penilai menyatakan bahawa model ini dinilai sebagai sangat tepat penggunaannya. Hal ini kerana model ini menggunakan teori binaan ayat BM terkini yang dianggap sangat membantu pengajaran dan pembelajaran ayat Bahasa Melayu. Selain itu, Puan Siti Salmah binti Sulaiman sebagai guru kanan yang juga ahli Munsyi Dewan menyatakan bahawa model ini dinilai sebagai sangat baik dan dapat membantu pengguna terutama murid sekolah dalam pemahaman struktur ayat BM. Namun, terdapat cadangan untuk penambahbaikan yang diberikan seperti dalam Jadual 4.5. Cadangan ini diperolehi hasil penilaian instrumen yang diberikan. Sebagai rumusan atau cadangan, instrumen soal selidik yang dikemukakan turut memberi ruang kepada penilai untuk mengemukakan cadangan jika berkaitan (rujuk Lampiran E).

Jadual 4.5

Cadangan Penambahbaikan Model VPS dengan Output Tambahan

Cadangan	Tindakan
Anak panah dalam paparan model dari atribut perkataan ke VPS perlu diganti supaya menampakkan proses "pemilihan".	Jenis garisan anak panah diganti daripada berbentuk lurus kepada bentuk bersambung.
Masukkan mesej ralat dalam cadangan ayat.	Mesej ralat ditambah dalam komponen cadangan ayat.
Atribut terjemahan boleh diganti dengan atribut Kata Tugas kerana kata tugas terlibat dalam pembentukan ayat.	Atribut terjemahan dikekalkan kerana atribut kelas kata telah merangkumi atribut kata tugas.

Terdapat tiga cadangan yang diperolehi daripada pakar penilai untuk penambahbaikan model. Jenis garisan anak panah yang menunjukkan VPS membuat paparan atribut perkataan diganti daripada penggunaan garisan lurus kepada garisan bersambung. Hal ini bertujuan untuk menunjukkan bahawa VPS hanya akan memaparkan atribut perkataan jika dipilih pautan oleh pengguna. Mesej ralat juga ditambah dalam bahagian semakan dan cadangan ayat kerana proses visualisasi juga akan memaparkan output mesej ralat jika tidak memenuhi skop ayat yang diperlukan atau jika terdapat ayat yang tidak dapat diproses. Penambahbaikan yang dilakukan dapat dirujuk dalam paparan model Rajah 4.6. Setelah penambahbaikan, model VPS dengan output tambahan ini digunakan untuk mereka bentuk algoritma berkaitan.



Rajah 4.6. Model VPS dengan output tambahan ditambah baik

4.2 Pembangunan Algoritma

Seperti yang telah dijelaskan dalam Bab Dua, kaedah asas pemrosesan bahasa tabii melibatkan kaedah **membahagikan ayat kepada perkataan**, membuat **padanan dengan rumus** atau kelas kata dan seterusnya menghasilkan **output**. Dalam kajian ini, kaedah sama dilalui dengan penambahan output mengikut skop kajian. Kaedah yang terlibat dibahagikan kepada tiga proses iaitu proses semakan, proses cadangan dan proses VPS. Proses semakan ayat melibatkan langkah 1) token ayat kepada perkataan, 2) semakan bilangan perkataan, 3) semak syarat ayat, 4) penandaan kelas kata, 5) semak

ejaan dan 6) semak rumus. Langkah seterusnya adalah untuk proses kedua iaitu 7) cadangan dan proses ketiga melibatkan turutan 8) VPS, 9) atribut perkataan, dan 10) VPS ayat contoh seperti yang ditunjukkan dalam Rajah 4.7.

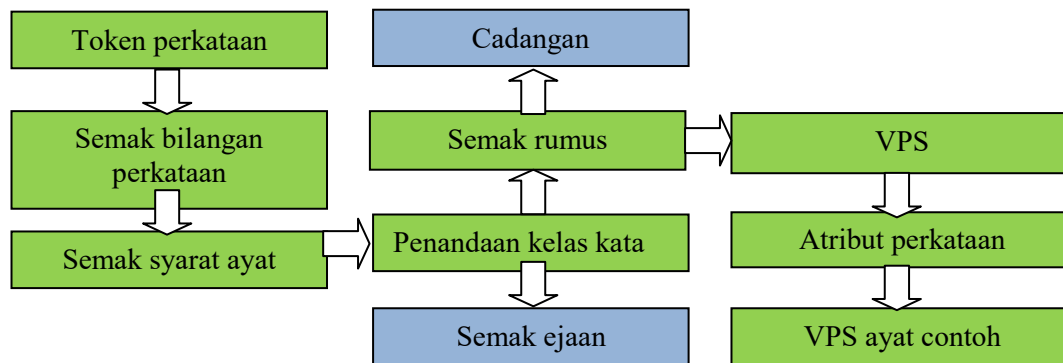
Komponen model VPS dengan output tambahan		Turutan pakej gabungan VPS	
1. Ayat	} rumus	} Teori <i>X-bar</i>	1. Token ayat
2. Frasa			2. Semak bilangan perkataan
3. Perkataan			3. Semak syarat ayat
4. Semakan			4. Penandaan kelas kata
5. Cadangan	} Teori Graf	} Teori Graf	5. Semak ejaan
6. VPS			6. Semak rumus
			7. Cadangan
			8. VPS
			9. Atribut perkataan
			10. VPS ayat contoh

Rajah 4.7. Perkaitan komponen model dan kaedah kajian.

Algoritma penyemak dengan cadangan pembetulan ayat dan VPS direka bentuk sebelum digabung menjadi algoritma VPS dengan output tambahan. Bahagian seterusnya menjelaskan turutan algoritma ini.

4.2.1 Algoritma VPS dengan Output Tambahan

Algoritma VPS dengan output tambahan melibatkan sepuluh langkah bermula dari penerimaan ayat input sehingga dapat menghasilkan VPS ayat contoh. Langkah tersebut adalah 1) token ayat kepada perkataan, 2) semakan bilangan perkataan, 3) semak syarat ayat, 4) penandaan kelas kata, 5) semak ejaan, 6) semak rumus, 7) cadangan, 8) VPS, 9) atribut perkataan dan 10) VPS ayat contoh seperti yang digambarkan dalam Rajah 4.8.



Rajah 4.8. Langkah algoritma VPS dengan output tambahan

Langkah satu hingga enam adalah algoritma semakan ayat. Manakala langkah ke tujuh adalah algoritma cadangan pembetulan ayat. Seterusnya langkah lapan hingga 10 adalah turutan algoritma VPS. Algoritma ini digabung dan dinamakan sebagai algoritma VPS dengan output tambahan. Algoritma ini diterjemahkan dalam bentuk persamaan matematik seperti berikut.

Turutan 1: Ayat dipisahkan (token) mengikut sempadan perkataan.

Turutan 2: Ayat (A) mesti tergolong dalam jumlah perkataan (P) yang melebihi atau sama dengan dua (2) perkataan dan tidak lebih atau sama 14 perkataan seperti persamaan berikut:

$$A = P_2 \leq P_n \leq P_{14} \quad (4.1)$$

Turutan 3: Semak syarat ayat dengan menapis ayat selain ayat penyata. Jenis ayat (JA) terdiri daripada ayat penyata (AP), ayat tanya (AT), ayat seruan (AS), ayat perintah (APe) dan ayat majmuk (AM). Untuk menyemak

syarat ayat (SA), semua perkataan kata nama (N), kata kerja (K), kata adjektif (A), kata sendi nama (KS) dan kata tugas (T) yang dipadankan dengan kelas kata (KK) tergolong dalam semua jenis ayat (JA) akan diterima kecuali (\setminus) ayat tanya, ayat seruan, ayat perintah dan ayat majmuk seperti persamaan berikut:

$$\begin{aligned} KK &= \{N, K, A, KS, T\} \\ JA &= \{AP, AT, AS, APe, AM\} \\ SA &= \{KK_5 \in JA \setminus AT, AS, APe, AM\} \end{aligned} \quad (4.2)$$

Turutan 4: Perkataan (P) yang mewakili bilangan perkataan dari dua hingga 14 mesti tergolong dalam jumlah keseluruhan KK yang ditetapkan untuk mendapat padanan yang sesuai seperti persamaan berikut:

$$P_i \in KK_5, i = 2, \dots, 14 \quad (4.3)$$

Turutan 5: Semak ejaan (SE) melibatkan proses menerima perkataan (P) dalam jumlah perkataan yang disyaratkan yang tidak tergolong dalam KK yang ditetapkan, namun tergolong dalam semua (\cap) format perkataan yang diterima yang dianggap sebagai kata nama (N) seperti persamaan berikut:

$$\begin{aligned} &\text{Format tarikh } \{dd/mm/yy\} \\ &\text{Format alamat } \{no, kg...00000...\} \\ &\text{Format kata nama khas } \{\text{'huruf besar'...}\} \\ N &= \{\text{alamat, tarikh, kata nama khas, nombor}\} \\ SE &= \{P_i \notin KK_5 \cap P_i \in N\}, i = 1, \dots, 14 \end{aligned} \quad (4.4)$$

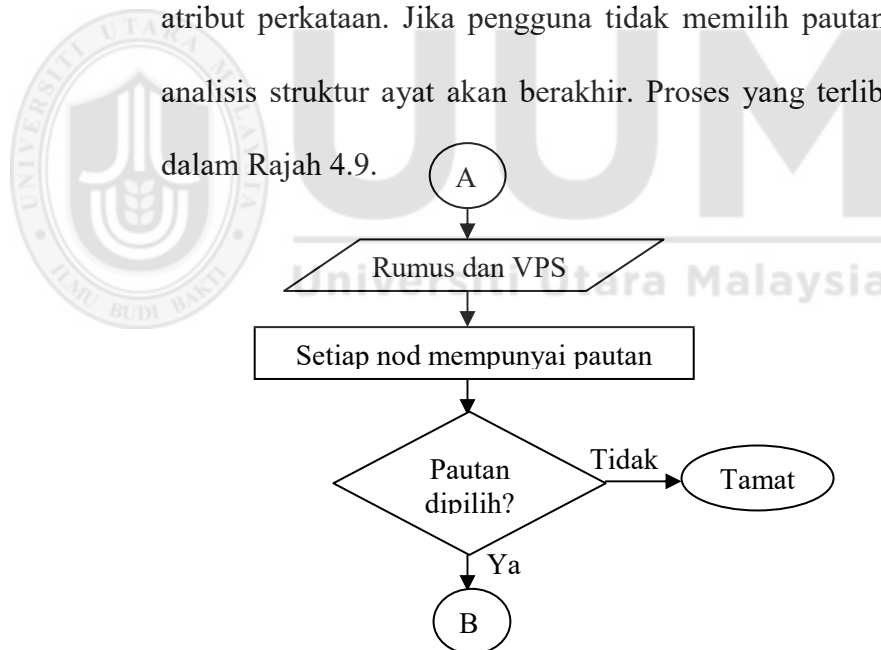
Turutan 6: Untuk membuat seakan rumus (SS), syarat ayat (SA) yang diterima mestilah tergolong dalam rumus yang ditetapkan seperti persamaan berikut:

$$SS = SA \in \text{Rumus} \quad (4.5)$$

Turutan 7: Ayat input (A) yang tidak sepadan, memerlukan pencarian rumus yang seakan sama dengan semua KK yang ditetapkan. Proses penggantian ayat dilakukan dengan berdasarkan kepada susunan KK.

$$A \in (KK_5 \cap \text{Rumus}) \quad (4.6)$$

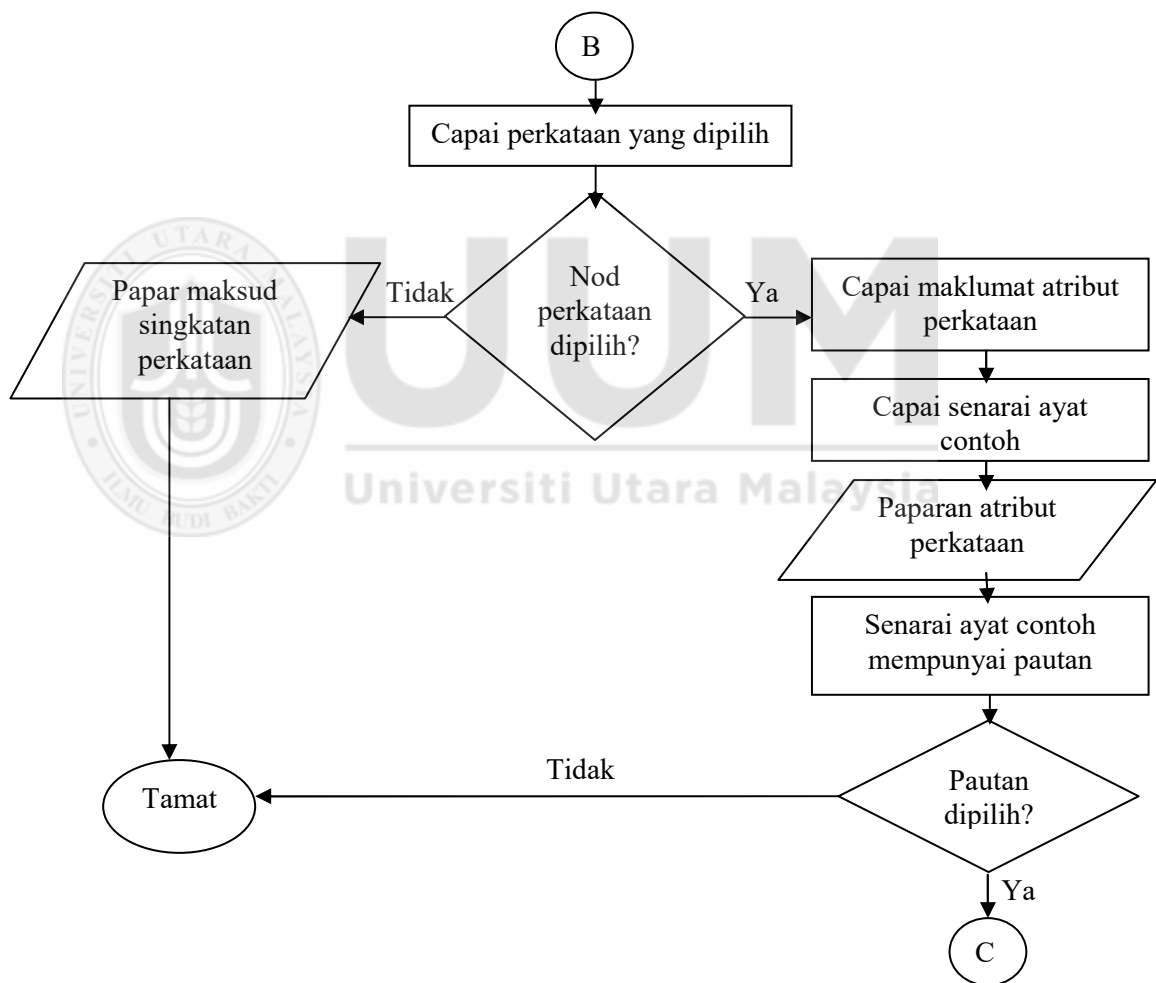
Turutan 8: Setiap nod dalam VPS boleh dipilih untuk membuat capaian paparan atribut perkataan. Jika pengguna tidak memilih pautan, maka proses analisis struktur ayat akan berakhir. Proses yang terlibat ditunjukkan dalam Rajah 4.9.



Rajah 4.9. Carta alir VPS

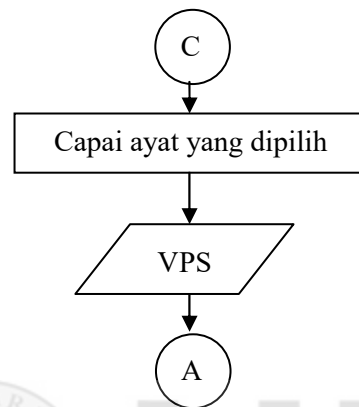
Turutan 9: Rajah 4.10 menunjukkan setelah pautan pada perkataan dalam VPS dipilih, atribut perkataan atau makna singkatan perkataan dipaparkan. Capaian maklumat dengan fail perkataan adalah untuk mendapatkan

maklumat berkaitan kelas kata, kata terbitan, dan terjemahan. Capaian dengan fail imej dibuat untuk mendapatkan imej berkaitan. Manakala capaian dengan fail ayat contoh adalah untuk mencapai senarai ayat yang terlibat. Setiap ayat contoh ini mempunyai pautan untuk membuat VPS yang baharu seperti yang dijelaskan dalam bahagian seterusnya.



Rajah 4.10. Carta alir atribut perkataan

Turunan 10: VPS ayat contoh adalah sama dengan VPS ayat input. Paparan atribut perkataan juga dipaparkan susunan ayat contoh yang lain berdasarkan kepada pemilihan perkataan yang dibuat. Proses yang terlibat ditunjukkan seperti dalam Rajah 4.11.



Rajah 4.11. Carta alir VPS ayat contoh

Jika pengguna tidak memilih pautan ayat contoh, maka VPS akan berakhir seperti yang ditunjukkan dalam paparan atribut perkataan Rajah 4.10. Jika pengguna memilih pautan, proses ulangan VPS akan dilakukan (Rajah 4.11). Algoritma berbentuk carta alir yang lengkap boleh dirujuk dalam Lampiran F.

4.3 Rumusan Bab Empat

Bab ini dimulakan dengan proses pembangunan model atribut perkataan yang menghasilkan model VPS dengan output tambahan. Komponen yang terdapat dalam model VPS adalah semakan, cadangan ayat, VPS, dan atribut perkataan. Penentuan model dilakukan oleh lima orang pakar bidang pengkomputeran linguistik dan Munsyi

Dewan. Sebanyak 10 soalan diberikan dalam soal selidik yang dibahagikan kepada kriteria penilaian umum dan kriteria penilaian komponen model. Kaedah yang dilakukan dalam menjalankan penentusahan ini adalah 1) memberi penerangan mengenai tujuan penentusahan, 2) memberi ruang kepada pakar menjalankan penilaian, 3) menganalisis dapatan, dan 4) membuat penambahbaikan model.

Seterusnya, algoritma penyemak dengan cadangan pembetulan ayat dan algoritma VPS direka bentuk. Algoritma tersebut menghasilkan algoritma VPS dengan output tambahan yang melibatkan sepuluh langkah. Antaranya 1) token ayat kepada perkataan, 2) semakan bilangan perkataan ($A = P_2 \leq P_n \leq P_{14}$), 3) semak syarat ayat ($SA = \{KK_5 \in JA \setminus AT, AS, APe, AM\}$), 4) penandaan kelas kata ($P_i \in KK_5, i = 2, \dots, 14$), 5) semak ejaan ($SE = \{P_i \in KK_5 \cap P_i \in N\}, i = 1, \dots, 14$), 6) semak rumus ($SS = SA \in Rumus$), 7) cadangan ($A \in (KK_5 \cap Rumus)$), 8) VPS, 9) atribut perkataan, dan 10) VPS ayat contoh. Algoritma ini digunakan untuk membina prototaip yang dijelaskan dalam Bab Lima seterusnya.

BAB LIMA

PEMBANGUNAN, LATIHAN, PENILAIAN PROTOTAIP DAN PERBINCANGAN

5.0 Pengenalan

Bab ini menjelaskan kaedah pembangunan prototaip berpandukan algoritma VPS dengan output tambahan. Prototaip ini diberi nama persona BMTutor. Prototaip yang dihasilkan memaparkan keunikan kajian dengan kejayaan menghasilkan output tambahan pohon sintaksis. Bab ini juga memberi fokus kepada fasa latihan dan penilaian prototaip untuk pembuktian konsep yang telah dikemukakan dalam model VPS dengan output tambahan. Hasil penilaian dianalisis mengikut kaedah penilaian dalam Parseval dan penilaian pengguna.

5.1 Aplikasi Teori Gestalt dan Teori Beban Kognitif Dalam Prototaip

Teori gestalt diaplikasikan dalam prototaip melalui kedudukan nod dan anak panah yang menggunakan prinsip kesinambungan. Prinsip ini mempunyai persamaan dengan teori graf untuk mereka bentuk anak panah dan nod. Anak panah yang digunakan perlu secara lurus dan tidak menyimpang sebagaimana yang telah dijelaskan dalam Bab Dua dan dijelaskan oleh Hicks (2009) bahawa paparan yang melibatkan nod dan anak panah, prinsip "*continuity*" atau kesinambungan boleh digunakan. Oleh yang demikian,

penggunaan anak panah lurus, bersambung antara nod dan tidak menyimpang diambil kira semasa reka bentuk VPS dilakukan.

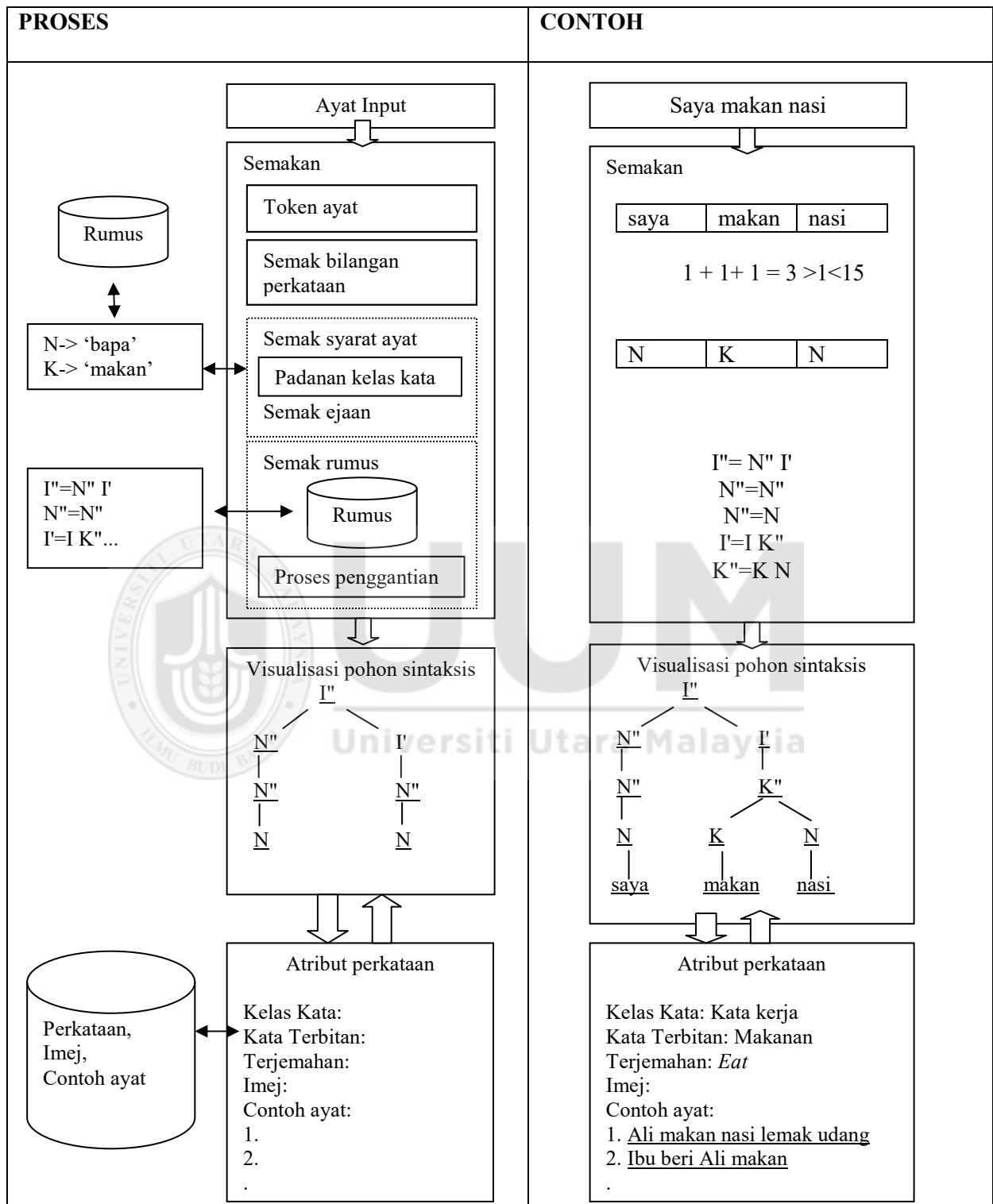
Prinsip lain seperti jarak, persamaan, penutupan dan kebiasaan juga diambil kira sebagaimana yang dicadangkan dalam Azizi *et al.* (2005). Jarak antara kedudukan nod dan anak panah menggunakan saiz yang sama iaitu dalam 10 sentimeter bagi meletakkan setiap nod agar membolehkan sesebuah perkataan dimuatkan. Persamaan pula menggunakan persamaan warna bagi membezakan warna nod dan anak panah. Selain itu, persamaan saiz perkataan juga dititikberatkan. Penutupan pula mencerminkan kedudukan anak panah dan nod yang digunakan adalah selari agar kedudukan perkataan, nod dan anak panah kelihatan menarik.

Teori beban kognitif diaplikasikan untuk tidak memuatkan banyak maklumat dalam skrin paparan. Paparan melibatkan kotak teks, butang dan VPS sahaja yang direka bentuk. Pembahagian skrin kepada dua bahagian untuk memaparkan VPS dan atribut perkataan adalah supaya paparan nampak lebih menarik dan tidak banyak melibatkan antara muka berbeza. Tahap kefahaman dan penerimaan yang diperolehi daripada penggunaan teori beban kognitif ini dinilai menggunakan soal selidik.

5.2 Reka Bentuk dan Pembangunan Prototaip

Lakaran antara muka prototaip dimulakan dengan mengumpul hasil paparan dari kajian terdahulu seperti dalam Lampiran A. Setiap penyambungan antara proses juga ditentukan berdasarkan kepada output yang dikeluarkan iaitu melibatkan kotak teks memasukkan input dan paparan output di bawahnya. Lakaran penuh berbentuk papan cerita dihasilkan yang menjadi panduan dalam membina kerangka pembangunan. Kerangka ini hanya melibatkan lakaran asas tentang pembangunan yang diperlukan. Lakaran sebenar direka bentuk menggunakan pengaturcaraan Python.

Pengaturcaraan Python digunakan untuk membina prototaip termasuk antara muka dan pangkalan data. Perisian ini dipilih berbanding bahasa pengaturcaraan lain seperti Prolog kerana perisian ini adalah perisian yang diakui oleh pembangun sistem NLP bahawa ia lebih banyak kelebihan berbanding perisian sealiran. Selain mudah diperolehi atau dimuat turun, perisian ini lebih mudah untuk mereka bentuk pohon sintaksis. Selain daripada itu, terdapat banyak kod pengaturcaraan untuk dirujuk yang disediakan dalam laman sesawang Python. Gambaran proses untuk menganalisis ayat dalam pembangunan prototaip ini ditunjukkan melalui Rajah 5.1.



Rajah 5.1. Proses menganalisis ayat

Dalam proses semakan (Rajah 5.1), proses semakan syarat ayat, semak ejaan dan semak rumus adalah proses pemilihan yang dilakukan oleh prototaip. Proses ini dilakukan berdasarkan syarat yang diberikan dalam kod atur cara mengikut skop kajian. Proses penggantian memainkan peranan jika ayat yang disemak semasa 'semak rumus' tidak sepadan. Justeru, proses penggantian melibatkan algoritma cadangan pembetulan ayat dilakukan dan proses seterusnya untuk membuat paparan VPS tidak dilakukan, sebaliknya mesej ralat berbentuk ayat yang betul sebagai cadangan kepada pengguna dipaparkan.

Terdapat aliran masuk dan keluar semasa paparan atribut perkataan kerana ayat contoh yang diberikan mempunyai pautan untuk membuat VPS yang baharu. Proses pepadanan dan rujukan terhadap pangkalan data juga menunjukkan aliran ulangan sehingga semua atribut perkataan dapat dipadankan. Selain itu, gambaran graf dilakukan mengikut jenis hierarki dan tidak bersilang berdasarkan penggunaan teori terlibat.

5.2.1 Menghubungkan Pangkalan Data dan Antara Muka

Pangkalan data yang dibangunkan terdiri daripada fail rumus, fail perkataan, fail imej, fail ayat contoh, fail contoh ayat, fail ayat majmuk dan fail istilah. Paparan antara muka dihubungkan dengan pangkalan data yang juga menggunakan pengaturcaraan Python.

Fail rumus (Rajah 5.2) menyimpan semua rumus dan kelas kata yang diperoleh hasil analisis yang dibuat ke atas ayat terkumpul. Format pengekodan Python digunakan dengan menggunakan simbol (->) sebagai pemisah antara perkataan terminal dan bukan terminal. Simbol (|) pula digunakan sebagai pemisah bagi setiap rumus yang membawa maksud 'atau'. Bagi menandakan kelas kata, setiap perkataan yang diwakili oleh kelas kata tersebut dipisahkan mengikut simbol (|) dan (' '). Penandaan ini boleh menyebabkan satu perkataan diwakili oleh lebih daripada satu kelas kata kerana faktor kekaburan pada perkataan tersebut yang bergantung kepada maksud dan jenis ayat yang digunakan. Proses penentuan ini ditentukan semasa fasa mengkategorikan ayat dilakukan (Bab Tiga).

Fail rumus (Rajah 5.2) digunakan semasa kaedah 'semak rumus' dilakukan. Pemadanan ini dilakukan semasa fasa semakan ayat dan proses memberi cadangan. Semasa proses semakan ayat, setiap perkataan dipadankan dengan kelas kata dan susunan rumus terlibat. Semasa memberi cadangan ayat, hasil jujukan frasa yang seakan sama dipadankan dengan semua perkataan ayat input dan proses pengubahan kedudukan perkataan dilakukan.

```

I" -> N" I'
N" -> N' Pent | N K | N | N A | N N' | N' N' | N' K" | N KS" | N N | N E
N' -> N N | KB N | Kbil N | KNafi N | N Pent | N N' | N' KS" | N K | N F
K" -> K | K N' | K' K' | K N | K A" | K KS" | K N" | N' N | K' N" | K KE
K' -> K A | Kad K | K KS" | K N | KB K | K K | K' N" | K N' | K' KS" |
KS" -> KS N' | KS N | KS N" | KS' N" | KS' K' | KS' K" | KS' KS'
KS' -> KS N" | KS N | KS N'
I' -> I A" | I N" | I K" | I' I' | I' K" | I I | I A | I KS"
A" -> A | KP A' | KB A | A N | A' KS" | KNafi A' | A N' | A K' | A' K" |
A' -> A K' | KP A | A N | A Kad | A K

```

Rajah 5.2. Keratan fail rumus.cfg.

Fail perkataan (Rajah 5.3) menyimpan empat komponen output atribut perkataan yang terdiri daripada terjemahan (T), kelas kata (GK), kata terbitan (KT) dan imej (I). Fail ini digunakan semasa padanan atribut perkataan. Komponen imej memerlukan fail perkataan untuk membuat pautan dengan fail imej. Manakala komponen ayat contoh pula diasingkan ke dalam fail yang berlainan (fail ayat contoh).

```

umur | T:age | GK:'Kata Nama' | KT:'berumur' | I:imej/umur.jpg
tahun | T:year | GK:'Kata Nama' | KT:'tahunan, bertahun' | I:i
masyarakat | T:society | GK:'Kata Nama' | KT:'bermasyarakat' |
melayu | T:malay | GK:'Kata Nama' | KT:'melayukan' | I:imej/me
kain | T:cloth | GK:'Kata Nama' | KT:'...' | I:imej/kain.jpg
songket | T:songket | GK:'Kata Nama' | KT:'...' | I:imej/songk
khairi | T:khairi | GK:'Kata Nama' | KT:'...' | I:imej/khairi.
laman | T:laman | GK:'Kata Nama' | KT:'lamankan, berlaman' | I:
kita | T:we | GK:'Kata Nama' | KT:'kitakan' | I:imej/kita.jpg
kad | T:card | GK:'Kata Nama' | KT:'...' | I:imej/kad.jpg

```

Rajah 5.3. Keratan fail perkataan.cfg

Rekod perkataan seperti dalam Rajah 5.3 di atas menunjukkan bahawa setiap perkataan akan dihubungkan dengan fail imej (Rajah 5.4) untuk mendapatkan imej yang berkaitan.

Fail ini adalah fail yang menyimpan imej bagi perkataan yang dikumpul. Imej yang dipaparkan hanya untuk perkataan dasar sahaja.



Rajah 5.4. Keratan fail imej

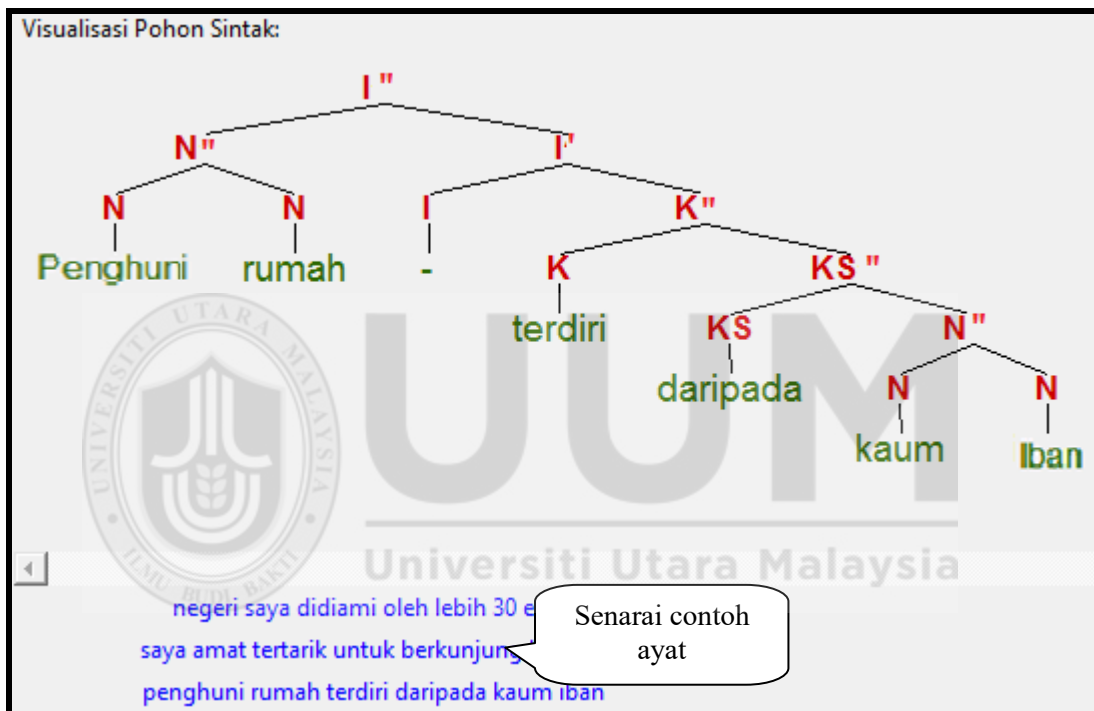
Fail ayat contoh (Rajah 5.5) adalah untuk menyimpan ayat contoh mengikut perkataan yang dikumpul. Perkataan yang dipilih semasa VPS akan mempunyai lebih dari satu konteks ayat sebagai panduan kepada pengguna tentang penggunaan ayat dalam konteks yang lain. Fail ini digunakan semasa padanan atribut perkataan dilakukan untuk memaparkan senarai ayat contoh.

```
makan : saya makan nasi | Dia makan dengan sabar | Ahmad suka ma
rumah : Rumah saya besar | Saya suka rumah saya
harganya : Dia letak harganya | Harganya boleh tahan | Harganya
saya : Saya anak baik | Saya pergi ke rumah awak | Saya sangat p
kevin : Kawan saya nama kevin | Kevin anak pak usu | Kevin belaj
menurut : saya menurut sahaja perintah awak | Dia menurut pak us
ketawa : Saya ketawa dengan kevin | ketawa menurut pak usu adala
masa : Dia perlukan masa secukupnya | Fikirkan masa yang suntuk
datuk : Dia datuk saya | Rosli bercakap dengan datuk itu
kemerdekaan : Esok kita sambut kemerdekaan | Dia rai kemerdekaan
bersalah : Dia rasa bersalah | Fikri bersalah dengan kevin
Luasnya : Kawasan itu luasnya memang besar | Rumah itu luasnya n
hadiah : dia beri sava hadiah | hadiah itu warna kuning | hadiah
```

Rajah 5.5. Keratan fail ayat contoh.cfg

Fail contoh ayat digunakan untuk menyimpan senarai ayat cadangan untuk kegunaan pengguna semasa memasukkan ayat input. Bagi pengguna yang tidak mempunyai

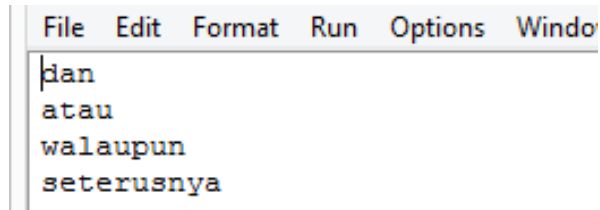
gambaran ayat yang hendak digunakan, senarai contoh ayat yang diberikan boleh dipilih seperti yang terdapat dalam paparan antara muka (Rajah 5.6). Pengguna boleh memilih untuk memasukkan ayat tersebut dalam kotak teks secara manual atau memilih pautan pada senarai ayat tersebut. Setiap ayat yang diberikan mempunyai pautan (*hyperlink*) secara terus untuk membuat VPS.



Rajah 5.6. Paparan senarai contoh ayat

Fail ayat majmuk menyimpan perkataan tapisan untuk menyemak syarat penerimaan ayat. Untuk membuat semakan ini, senarai perkataan yang berkaitan dengan ayat tapisan disenaraikan seperti contoh yang ditunjukkan dalam Rajah 5.7. Jika semakan sepadan

dengan salah satu perkataan tapisan, maka mesej ralat dikeluarkan. Proses ini dilakukan selepas kaedah semak bilangan perkataan berjaya dilakukan.



```
File Edit Format Run Options Window
| dan
| atau
| walaupun
| seterusnya
```

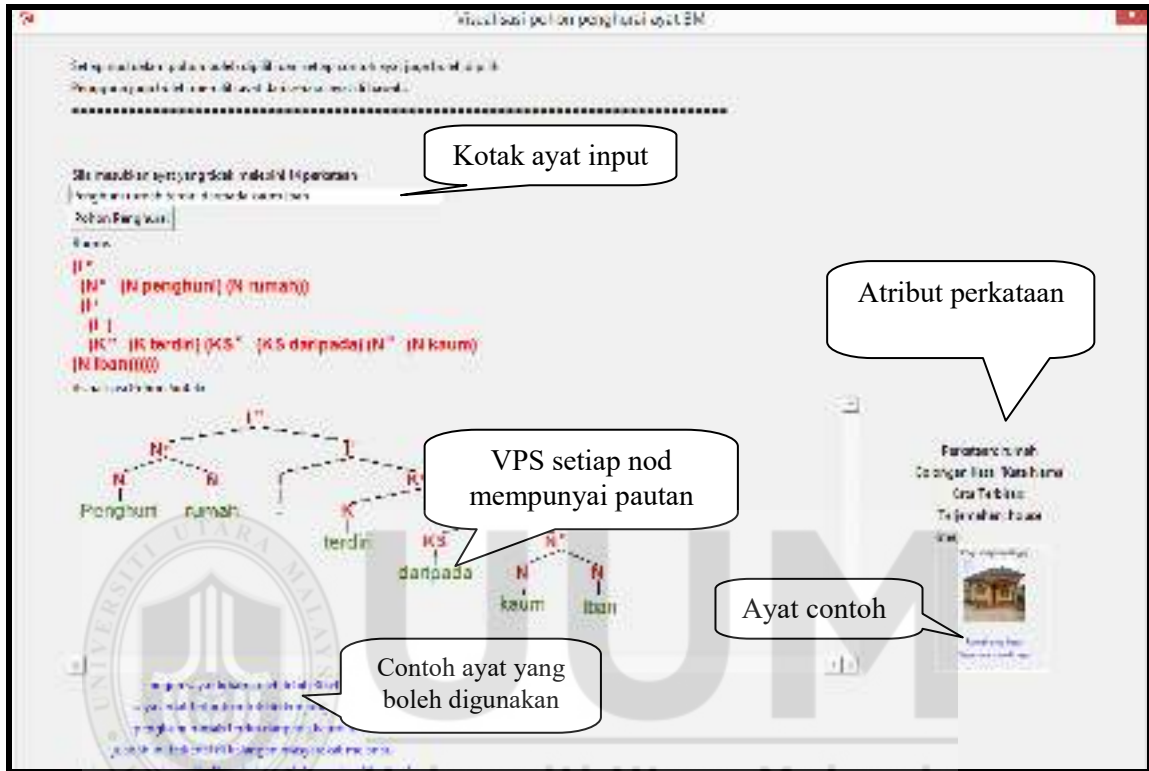
Rajah 5.7. Keratan fail ayat majmuk.cfg

Setelah semua fail pangkalan data yang diperlukan direka bentuk, pengkodan atur cara untuk menghubungkan antara muka prototaip dan pangkalan data dilakukan. Seterusnya, pengkodan atur cara pembangunan VPS dilakukan yang dijelaskan dalam bahagian seterusnya.

5.2.2 Pengkodan Atur Cara Pembangunan VPS

Paparan antara muka dibahagikan kepada dua bahagian untuk paparan VPS dan atribut perkataan dalam satu paparan skrin. Proses sebelum penyambungan dibuat kepada pangkalan data, adalah pembangunan antara muka secara asas dan setelah pangkalan data siap dibangunkan proses menambahbaik antara muka dijalankan. Proses ini lebih mencabar kerana VPS juga perlu dihasilkan daripada ayat contoh yang disediakan dalam paparan atribut perkataan. Selain itu, paparan VPS ayat contoh perlu diberi pautan pada setiap perkataan untuk menghasilkan atribut perkataan. Strategi lain diperlukan apabila paparan skrin perlu dibahagikan kepada dua bahagian untuk memaparkan output VPS

dan atribut perkataan supaya output lebih mesra pengguna. Rajah 5.8 berikut menunjukkan contoh paparan antara muka utama.



Rajah 5.8. Antara muka BMTutor

Di sebalik paparan VPS yang dapat dihasilkan, proses lain sebelumnya akan dikeluarkan berdasarkan kepada algoritma yang direka bentuk. Setiap proses yang terlibat dalam memaparkan paparan VPS yang berjaya seperti di atas ditunjukkan dalam bahagian seterusnya mengikut padanan turutan algoritma VPS dengan output tambahan.

5.3 Antara Muka Prototaip

BMTutor menghasilkan paparan VPS berserta rumus binaan ayat dan atribut perkataan. Disebalik proses memaparkan output tersebut, terdapat proses lain yang terlibat seperti semakan dan cadangan. Paparan output ini ditunjukkan dalam bahagian 5.3.1 hingga 5.3.4 mengikut turutan algoritma.

5.3.1 Token Perkataan dan Semak Bilangan Perkataan

Ayat input dibuat pembahagian perkataan (token) untuk proses mengira jumlah perkataan. Rajah 5.9 berikut memberi contoh paparan yang dikeluarkan jika pengguna tidak memasukkan ayat mengikut syarat algoritma $A = P_2 \leq P_n \leq P_{14}$ iaitu kurang dari dua perkataan dan lebih dari 14 perkataan.

Sila masukkan ayat yang tidak melebihi 14 perkataan

Saya

Pohon Penghurai

Bilangan perkataan mestilah antara 2 hingga 14

Sila masukkan ayat yang tidak melebihi 14 perkataan

yam nasi kuah nasi ayam nasi putih nasi berlauk nasi apa nasi sedap

Pohon Penghurai

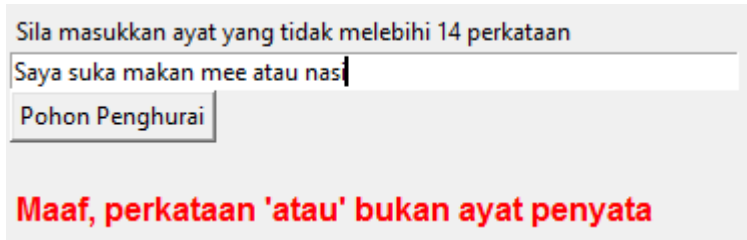
Ayat anda melebihi 14 perkataan

Rajah 5.9. Semak bilangan perkataan

Jika syarat jumlah perkataan dipenuhi, proses semakan ayat, kelas kata dan ejaan dilakukan.

5.3.2 Semak Syarat Ayat, Penandaan Kelas Kata, Semak Ejaan

Bagi ayat yang dikenal pasti bukan ayat penyata, mesej ralat seperti dalam Rajah 5.10 dipaparkan yang menyatakan perkataan tersebut mewakili ayat yang tidak dianalisa.



Sila masukkan ayat yang tidak melebihi 14 perkataan

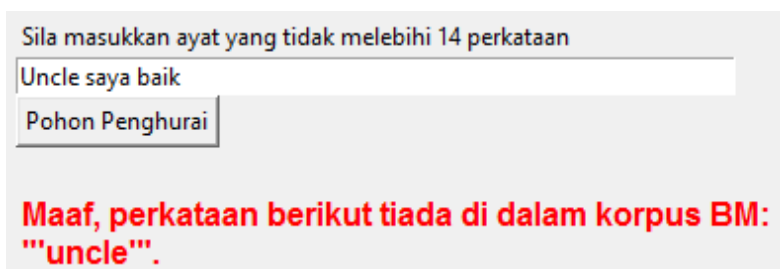
Saya suka makan mee atau nasi

Pohon Penghurai

Maaf, perkataan 'atau' bukan ayat penyata

Rajah 5.10. Semak syarat ayat

Seterusnya, jika syarat ayat masih dipenuhi mengikut skop yang ditetapkan, setiap perkataan dalam ayat input dipadankan dengan kelas kata yang berkaitan. Namun jika terdapat perkataan yang tidak dapat dipadankan, mesej ralat dipaparkan. Bagi mengesahkan bahawa perkataan tersebut bukan kata nama khas, nombor, tarikh atau alamat, maka semakan ejaan dipadankan dengan format yang ditetapkan. Jika tidak memenuhi syarat, maka mesej ralat seperti Rajah 5.11 dipaparkan.



Sila masukkan ayat yang tidak melebihi 14 perkataan

Uncle saya baik

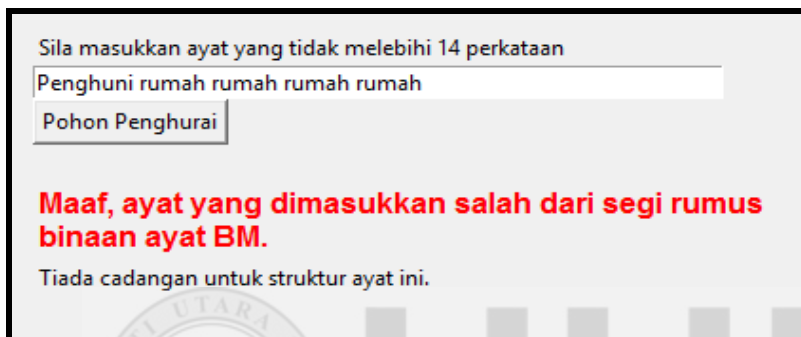
Pohon Penghurai

Maaf, perkataan berikut tiada di dalam korpus BM: "uncle".

Rajah 5.11. Perkataan yang tiada dalam simpanan

5.3.3 Semak Rumus, Cadangan, VPS

Semasa semakan rumus, jika ayat yang dimasukkan tidak sepadan, maka mesej ralat dipaparkan. Bagi ayat yang tidak mempunyai susunan rumus yang seakan sama, maka mesej ralat seperti dalam Rajah 5.12 dipaparkan menyatakan bahawa tiada cadangan untuk ayat tersebut.



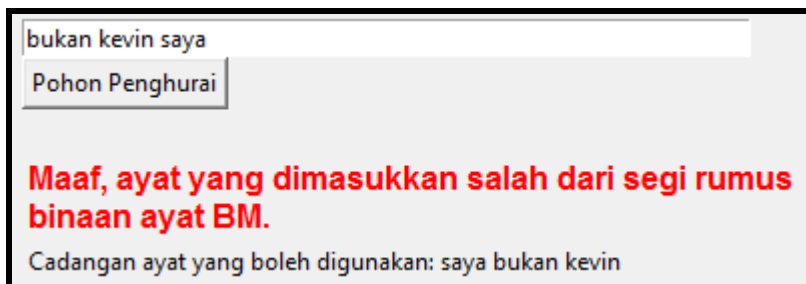
Sila masukkan ayat yang tidak melebihi 14 perkataan
Penghuni rumah rumah rumah rumah
Pohon Penghurai

Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.

Tiada cadangan untuk struktur ayat ini.

Rajah 5.12. Ayat yang tidak dapat diproses

Semasa semakan juga, jika ayat didapati salah dari segi penggunaan rumus tetapi hasil padanan mendapati terdapat susunan rumus yang seakan sama, maka mesej ralat dikeluarkan berserta dengan cadangan pembedahan ayat seperti dalam Rajah 5.13.



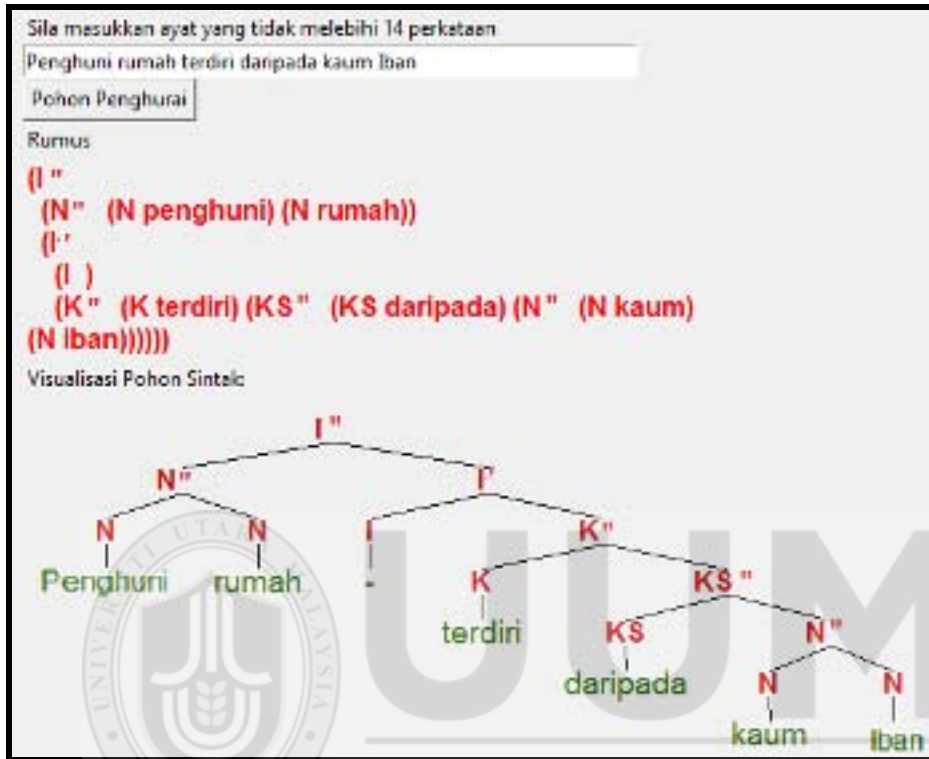
bukan kevin saya
Pohon Penghurai

Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.

Cadangan ayat yang boleh digunakan: saya bukan kevin

Rajah 5.13. Cadangan pembedahan ayat

Jika ayat yang dimasukkan menepati semua syarat dan padanan, maka output VPS dipaparkan berserta dengan susunan rumus seperti yang ditunjukkan dalam Rajah 5.14.



Rajah 5.14. VPS ayat input

Daripada VPS seperti dalam Rajah 5.15, setiap nod yang mewakili perkataan dalam ayat boleh dipilih untuk mengetahui atribut perkataan yang terlibat. Pautan (*hyperlink*) disertakan pada setiap perkataan. Jika pautan dipilih (klik), maka paparan atribut dipaparkan.

5.3.4 Atribut Perkataan

Setiap nod dalam VPS boleh dipilih kerana mempunyai pautan. Jika nod perkataan bagi ayat input dipilih contohnya perkataan "rumah", maka atribut perkataan seperti dalam Rajah 5.15 dipaparkan.

(a) Perkataan: rumah
Golongan Kata: 'Kata Nama'
Kata Terbitan:
Terjemahan: house
Imej:
Imej: imej/numah.jpg
Rumah saya besar
Saya suka rumah saya

(b) Sila masukkan ayat yang tidak melebihi 14 perkataan
Rumah saya besar
Pohon Penghurai
Rumus
(I'' (N'' (N rumah) (N saya)) (I'' (I'' (A'' (A besar))))
Visualisasi Pohon Sintak:
I''
N'' I''
N N I A''
Rumah saya - A
besar

Rajah 5.15. Atribut perkataan dan VPS ayat contoh

Senarai ayat contoh dalam paparan atribut perkataan seperti dalam Rajah 5.16 paparan (a) boleh dipilih kerana mempunyai pautan (*hyperlink*). Jika salah satu ayat dipilih, VPS yang baharu dipaparkan seperti dalam paparan (b) iaitu bagi ayat "Rumah saya besar". Sebelum paparan (b) dapat dilakukan, turutan proses yang sama (ulangan) akan berlaku sebagaimana proses permulaan iaitu bermula dengan kaedah token perkataan.

5.4 Output Tambahan

Berbanding kajian terdahulu, kajian ini telah menambah empat output lain seperti 1) rumus BM *X-bar*, 2) penyemak dengan cadangan pembetulan ayat, 3) atribut perkataan, dan 4) VPS ayat contoh. Setiap output tambahan ini dijelaskan dalam bahagian seterusnya.

5.4.1 Rumus *X-bar*

Nik Safiah *et al.* (2009) telah menyediakan rumus asas binaan ayat seperti yang ditulis dalam buku Tatabahasa Dewan Edisi Ketiga. Rumus tersebut mewakili ayat dasar BM yang ditambah baik daripada rumus yang dikeluarkan dalam buku Nik Safiah, Farid, Hashim dan Abdul Hamid (2004). Kajian yang dilakukan oleh Ahmad Izuddin *et al.* (2007) adalah merujuk kepada rumus yang dikeluarkan dalam buku Nik Safiah (1995) dan Nik Safiah *et al.* (2004). Manakala Noor Hafhizah (2011) pula merujuk kepada buku Nik Safiah (1995). Rumus binaan ayat tersebut dinamakan sebagai RSF. Oleh kerana rujukan rumus para pengkaji sebelumnya adalah merujuk RSF, maka rumus BM *X-bar* yang diperkenalkan dalam kajian ini menjadi satu keunikan dalam penggunaan VPS kerana rumus *X-bar* adalah hasil peningkatan rumus oleh pakar bahasa terhadap kelemahan RSF. Selain itu, rumus *X-bar* diakui oleh para pengkaji bahasa bahawa rumus ini lebih universal dengan penggunaan rumus yang lebih meluas dan kebarangkalian untuk memaparkan output kabur struktur dalam kajian pohon sintaksis adalah rendah.

5.4.2 Cadangan Pembetulan Ayat

Cadangan pembetulan ayat berdasarkan rumus *X-bar* boleh meningkatkan ketepatan output VPS. Cadangan ini juga boleh membantu pengguna memahami penggunaan ayat yang gramatis selain memberi kepentingan kepada bidang pemprosesan bahasa.

5.4.3 Atribut Perkataan

Pautan pada nod VPS untuk menghasilkan atribut perkataan memberi idea dan pemahaman lebih mendalam tentang penggunaan sesebuah perkataan. Atribut ini juga boleh membantu pengguna untuk mendapatkan atribut lain. Sebagai contoh, atribut imej yang diberikan boleh menterjemahkan situasi berbeza sebagai contoh, imej "makan" membawa maksud "makanan/dimakan/termakan" yang membolehkan pengguna berfikir dengan kreatif.

5.4.4 VPS Ayat Contoh

Senarai ayat contoh yang diberikan bukan hanya dapat memberi cadangan ayat dalam konteks yang lain tetapi boleh membantu pemahaman tentang ayat berbeza. Ayat ini boleh membantu pengguna untuk membezakan penggunaan ayat bagi perkataan yang sama.

Output tambahan yang dijelaskan dalam bahagian 5.4.1 hingga 5.4.4 telah memberi empat sumbangan dalam kajian ini. Sumbangan ini dijelaskan dalam Bab Satu iaitu 1) model VPS berserta output tambahan, 2) algoritma VPS dengan semakan dengan cadangan pembetulan ayat, 3) rumus $X\text{-bar}$ bagi kegunaan VPS, dan 4) pengecaman atribut perkataan.

5.5 Latihan Prototaip

Bagi memastikan ketepatan output diperolehi adalah tinggi, maka persoalan kedua kajian "Bagaimana cara untuk memastikan model dan algoritma yang dihasilkan adalah tepat?" perlu di jawab. Salah satu kaedah yang digunakan adalah dengan menjalankan proses latihan prototaip. Proses latihan ini bagi memastikan kekangan rumus yang boleh ditambah baik. Ini bertujuan untuk menghasilkan VPS yang tidak hanya dapat memaparkan output malah dapat memaparkan output VPS dengan tepat mengikut struktur pembentukan ayat BM. Sebanyak 10% (106 ayat) daripada ayat terkumpul untuk fasa latihan prototaip (1055 ayat) diambil secara rawak untuk menjalankan setiap proses latihan seperti yang ditunjukkan dalam Jadual 5.1.

Jadual 5.1

Jumlah Ayat Untuk Latihan Prototaip

Pola Ayat	Peratusan	Latihan	10% ayat
FN + FN	12%	127	13
FN + FK	65%	686	69

Sambungan Jadual 5.1

Pola Ayat	Peratusan	Latihan	10% ayat
FN + FA	16%	169	17
FN + FSN	7%	74	7
Jumlah	100%	1055	106

Latihan pertama yang dilakukan berjaya menghasilkan 61 VPS tepat. Manakala selebihnya 45 VPS didapati tidak tepat kerana menghasilkan VPS melebihi daripada satu. Penghasilan VPS ini disebabkan penggunaan rumus yang boleh ditambah baik contohnya $K N$ dan $K N''$ akan mengeluarkan padanan dua output VPS. Hal ini kerana kelas kata N disimpan juga dalam frasa nama (N'').

Rumus kemudiannya ditambah baik dengan hanya menyimpan rumus kelas kata N dalam susunan frasa nama (N''). Biasanya kelas kata yang diwakili oleh frasa dalam kekangan ini adalah kelas kata di penghujung rumus untuk membentuk nod pohon yang baharu. Justeru itu, kelas kata yang telah diwakili oleh frasa telah menyimpan kelas kata yang sama.

Latihan kedua telah mengambil sebanyak 106 ayat yang berbeza yang dikategorikan untuk fasa latihan. Hasil yang diperoleh adalah sebanyak 82 VPS tepat. Selain daripada itu, hasil latihan juga menunjukkan prototaip berjaya membuat VPS bagi semua ayat

yang dimasukkan. Hal ini menunjukkan rumus $X\text{-bar}$ boleh digunakan untuk VPS.

Jadual 5.2 menunjukkan ringkasan hasil latihan prototaip yang dilakukan.

Jadual 5.2

Hasil Uji Kaji Fasa Latihan Prototaip Secara Keseluruhan

Latihan (L)	L 1	L2
Jumlah ayat	106	106
VPS	106	106
VPS tepat	61	82
VPS melebihi satu output	45	24
Kekangan	Penggunaan rumus yang boleh ditambah baik contohnya $K N$ dan $K N''$ akan mengeluarkan padanan dua output VPS	
Penambahbaikan	Membuang kelas kata yang telah diwakili oleh frasa. Contohnya rumus $K N$ dibuang disebabkan telah terdapat rumus $K N''$ yang juga akan membuat padanan N'' bersama dengan N	

Oleh yang demikian, secara keseluruhan, rumus yang boleh digunakan untuk tujuan VPS ayat penyata BM boleh mengandungi rumus seperti dalam Jadual 5.3. Rumus penuh disediakan dalam Lampiran G.

Jadual 5.3

Rumus X-bar

I' -> N' I'

I' -> A" | I A" | I N" | I K" | I' I' | I' K | I I | I A | I KS" | K"

SN" = N + (N) + (penerang) + (*KBIL) + (*PENT) + (KS") + (*KAD) + (*KNF) + (*KPM) + (*KB)

N" = N + (N) + (KB) + (*PENT) + (A") + (K") + (*KBIL) + (KS") + (*KAD) + (*KNF) + (*KPM) + (KARAH)

K" = (KPN) + (KPM) + K + (*KB) + (*KAD) + (N") + (A") + (*KNF) + (KP) + (KS")

A" = (KP) + (KPM) + (KNF) + A + (A) + (*KAD) + (*KB) + (N") + (K") + (KS")

KS" = (KPM) + (KB) + (KAD) + (KNF) + KS + N" + (KS)

*Boleh wujud di mana-mana sama ada dihadapan/belakang kelas kata yang lain dan boleh wujud melebihi dari satu posisi iaitu gabungan boleh terdiri daripada satu atau melebihi dari satu dalam satu ayat sama ada secara berderetan atau pun tidak

Jadual 5.3 menunjukkan frasa nama subjek (SN") boleh mempunyai unsur penerang berupa penerang kata kerja atau penerang kata adjektif. Struktur frasa nama (N"), frasa kerja (K") dan frasa adjektif (A") boleh berdiri sendiri atau digabung dengan kelas kata yang lain. Frasa nama (N") juga diperlukan untuk digabung bersama kata sendi nama (KS) bagi membentuk frasa sendi nama (KS"). Setiap frasa akan mempunyai frasa pertengahan sebagai contoh N" mempunyai frasa pertengahan N'. Rumus ini digunakan untuk menjalankan fasa penilaian prototaip dalam bahagian seterusnya.

5.6 Penilaian Prototaip

Penilaian prototaip dilakukan untuk mencapai objektif ketiga iaitu untuk "Menilai output VPS dan cadangan pembetulan ayat berdasarkan metrik penilaian dalam pemrosesan bahasa tabii untuk menguji ketepatan output dan membuat penilaian pengguna bagi mencapai kepuasan subjektif dan penerimaan kognitif pengguna". Oleh kerana model VPS dengan output tambahan mempunyai empat komponen iaitu semakan, cadangan, VPS, dan atribut perkataan, maka penilaian juga mesti meliputi komponen ini. Penilaian yang dilakukan dijelaskan dalam bahagian seterusnya.

5.6.1 Penilaian Parseval

Hasil uji kaji ditunjukkan dalam empat pola ayat iaitu ayat berpola frasa nama, frasa kerja, frasa adjektif dan frasa sendi nama seperti yang dilakukan oleh Noor Hafhizah (2011). Hasil uji kaji ke atas setiap pola ayat dikira berdasarkan ukuran metrik *recall* dan *precision*.

Penilaian output dalam kajian Ahmad Izuddin *et al.* (2007) mendapat peratusan 81.3% *recall* dan kajian Noor Hafhizah (2011) mendapat 100% *recall* dan 93.2% *precision*. Oleh yang demikian, kaedah menjalankan uji kaji yang digunakan dalam kedua-dua kajian tersebut dirujuk agar kajian ini dapat mencapai peratusan *recall* dan *precision* yang tinggi. Secara khusus, kajian ini akan menggunakan prototaip yang dibangunkan

untuk menentukan apakah model dan algoritma yang direka bentuk dapat melakukan penilaian dengan output yang tepat.

5.6.2 Ayat Uji Kaji

Jumlah ayat yang diambil untuk fasa uji kaji adalah sebanyak 151 ayat. Daripada jumlah ayat ini, dapat dirumuskan setiap pola ayat mengandungi jumlah bilangan ayat seperti dalam Jadual 5.4 (rujuk Lampiran H).

Jadual 5.4

Bilangan Ayat Uji Kaji Mengikut Pola Ayat BM

Pola ayat	Bilangan ayat terlibat
Frasa nama	18
Frasa kerja	98
Frasa adjektif	24
Frasa sendi nama	11
Jumlah ayat uji kaji	151

5.6.3 Hasil Uji Kaji VPS

Hasil uji kaji bagi setiap ayat menunjukkan prototaip berjaya membuat VPS secara tepat dengan satu output sebanyak 147 ayat daripada 151 ayat uji kaji. Sebanyak empat ayat menghasilkan dua output. Hasil uji kaji ini ditunjukkan seperti dalam Jadual 5.5.

Jadual 5.5

Ringkasan Hasil Uji Kaji VPS

Pola ayat	Bilangan ayat terlibat	Hasil VPS	Hasil VPS tepat
Frasa nama	18	18	17
Frasa kerja	98	98	95
Frasa adjektif	24	24	24
Frasa sendi nama	11	11	11
Jumlah ayat	151	151	147

Setiap pola ayat dinilai mengikut kategori bagi kedua-dua metrik penilaian yang ditunjukkan seperti dalam Jadual 5.6.

Jadual 5.6

Purata dan Peratus Hasil Uji Kaji VPS

Pola ayat	Recall Jumlah output tepat /Output tepat yang dikira	Precision Output tepat/Semua ayat uji kaji
Frasa nama	17/17=100%	17/18=94.4%
Frasa kerja	95/95=100%	95/98=96.9%
Frasa adjektif	24/24=100%	24/24=100%
Frasa sendi nama	11/11=100%	11/11=100%
Purata	100%	97.8%

Hasil ketepatan VPS mencapai purata 97.8% dan dapat mengeluarkan output 100% *recall*. Purata peratusan *recall* 100% yang diperolehi dalam kajian ini berjaya

mengeluarkan output untuk semua ayat uji kaji. Kajian Noor Hafhizah (2011) pula tidak dapat memaparkan kesemua output kerana hanya berjaya mengeluarkan 19 output daripada semua 26 ayat uji kaji.

Hasil uji kaji juga menunjukkan sebanyak empat ayat mengeluarkan dua output. Salah satu output yang dihasilkan adalah betul, manakala salah satu output lagi dikategorikan salah apabila padanan rumus mengambil hanya satu kata nama sebagai subjek bagi ayat yang mempunyai subjek melebihi dari satu kelas kata. Ayat berpola Frasa Kerja (K") mengeluarkan tiga output yang salah, dan ayat berpola Frasa Nama (N") juga mengeluarkan satu output yang sama seperti yang ditunjukkan dalam Jadual 5.7.

Jadual 5.7

Ayat Output Melebihi Satu

Ayat Uji kaji

1. Temenggung Abu_Bakar akhirnya diiktiraf sebagai Sultan johor
 2. Pelajar Islam mengambil kesempatan bersolat zuhur di masjid itu
 3. Ayah Khairi berbincang dengan keluarganya.
 4. Surat tidak rasmi disebut juga sebagai surat kiriman biasa.
-

Ayat seperti dalam Jadual 5.7 di atas mengeluarkan output melebihi satu disebabkan output yang dihasilkan menunjukkan VPS mengambil hanya satu kata nama mewakili subjek. Sebagai contoh, kata nama subjek bagi ayat "Ayah Khairi" telah membuat padanan dengan hanya mengambil kata nama "Ayah" sebagai subjek dan perkataan

"Khairi" dipisahkan menjadi predikat. Kesalahan padanan ini berpunca daripada padanan rumus yang semenangnya mengandungi hanya satu kata nama sebagai subjek atau melebihi dari satu. Oleh yang demikian, VPS yang dihasilkan adalah melebihi daripada satu berdasarkan padanan yang dibuat. Kesalahan padanan ini lebih mudah terjadi bagi ayat berpola frasa kerja (K").

Bagi ayat "Surat tidak rasmi disebut juga sebagai surat kiriman biasa", pula hanya mengambil kata nama "surat" sebagai subjek, manakala selebihnya "tidak rasmi disebut juga sebagai surat kiriman biasa" dipisahkan menjadi predikat. Padanan ini menghasilkan dua output VPS. Sama seperti ayat "Temenggung Abu_Bakar akhirnya diiktiraf sebagai Sultan johor", yang mengeluarkan salah satu output dengan subjek "Temenggung" dan selebihnya dipisahkan menjadi "Abu_Bakar akhirnya diiktiraf sebagai Sultan johor".

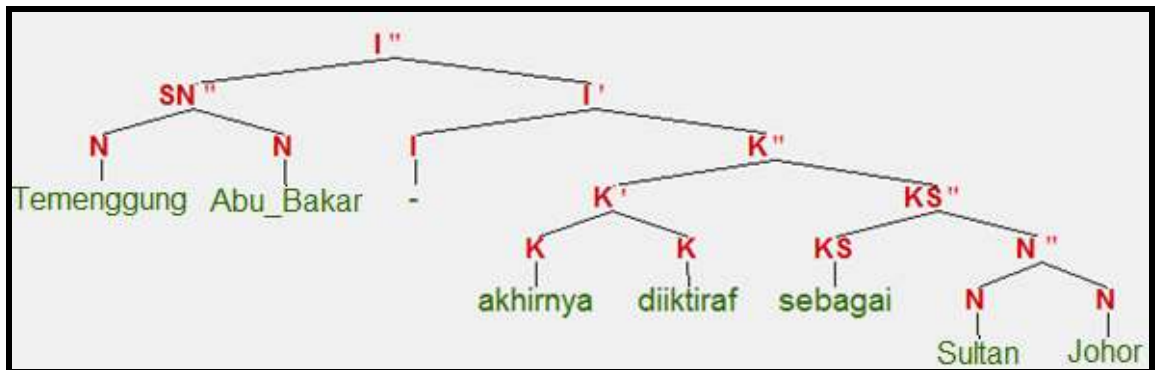
Oleh itu, pengasingan frasa nama untuk subjek yang mempunyai unsur penerang perlu dilakukan. Walaupun frasa nama ini diasingkan, masalah unsur penerang dalam subjek masih tidak dapat ditangani jika tidak diberikan kata yang unik. Hal ini kerana terdapat ayat yang mengandungi unsur penerang dalam subjek dan unsur tersebut tidak dapat diasingkan atau dibuang kerana akan menyebabkan kesalahan pada ayat asal. Antara contoh ayat tersebut adalah seperti dalam Jadual 5.8.

Jadual 5.8

Contoh Ayat dengan Penggunaan Unsur Penerang dalam Subjek

Ayat	Keterangan
S=Berita buruk itu P=belum disampaikan kepada kaum keluarga mereka.	Perkataan 'buruk' adalah unsur penerang kata adjektif kepada kata nama 'berita'
S=Kelebihan ketara penggunaan teknologi itu P=ialah teknologi tersebut amat berkesan untuk menyembuhkan penyakit berkenaan.	Perkataan 'ketara' adalah penerang kata adjektif kepada kata nama kelebihan
S=Hadiah tersebut P=cantik	Perkataan 'tersebut' adalah penerang kata kerja kepada kata nama 'hadiah'.

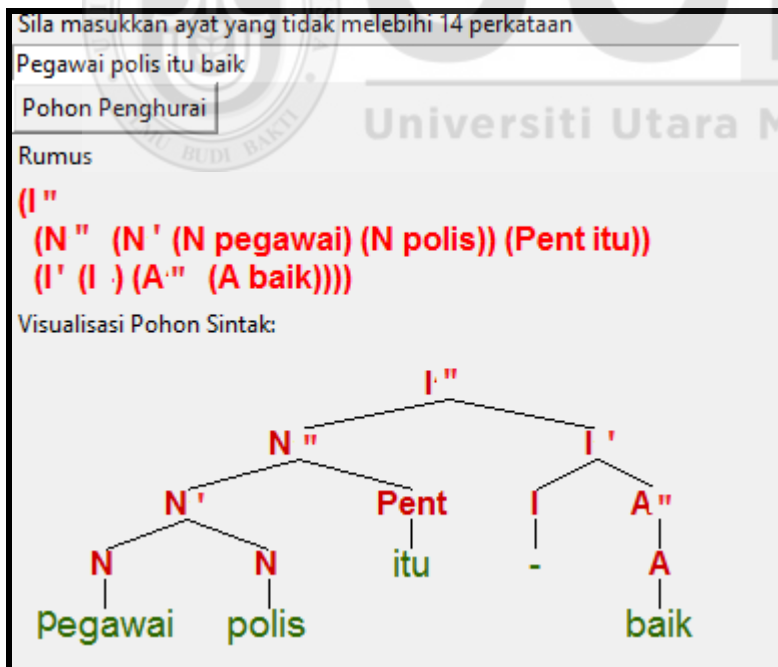
Penerang bagi kata kerja (PK) dan penerang kata adjektif (PA) dinamakan dalam frasa nama subjek (SN^{''}) sebagai penambahbaikan rumus. Perkataan yang termasuk dalam kelas kata ini akan mengeluarkan nod bagi SN^{''}. Namun, bagi kata yang tidak berunsur penerang, kelas kata dikekalkan sebagai contoh kelas kata nama (N). Rajah 5.16 berikut memberi contoh output bagi frasa nama subjek (SN^{''}).



Rajah 5.16. Contoh output frasa nama yang diasingkan

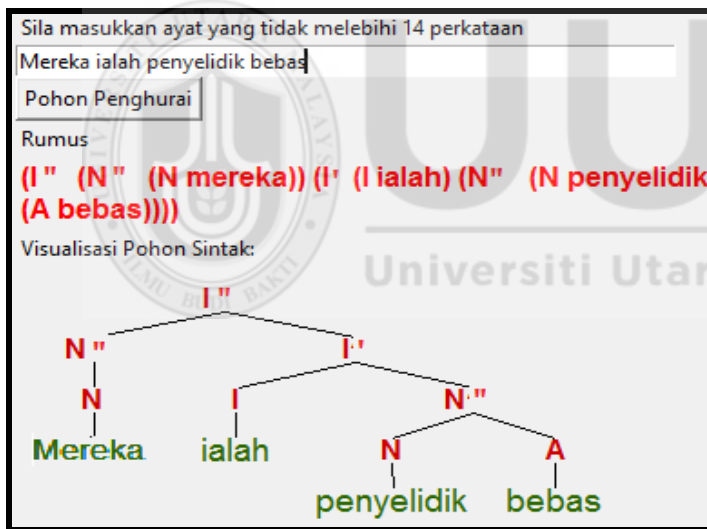
Setelah penambahbaikan dilakukan, tidak hanya VPS tepat dapat ditingkatkan, malah masalah output melebihi satu yang disebabkan unsur penerang dapat dielakkan. Selain daripada itu, prototaip berjaya membuat VPS bagi semua ayat yang dimasukkan. Rumus ditambahbaik dengan mengasingkan frasa nama untuk subjek dan penggunaan unsur penerang. Setiap frasa akan mempunyai frasa pertengahan sebagai contoh N'' mempunyai frasa pertengahan N'.

Daripada hasil VPS yang dipaparkan, bagi ayat input yang mempunyai jumlah perkataan kurang daripada tujuh akan menghasilkan output dengan cepat. Hal ini terjadi kerana bilangan perkataan sedikit memudahkan pencarian padanan dibuat. Rajah 5.17 hingga Rajah 5.17 adalah contoh ayat ini.



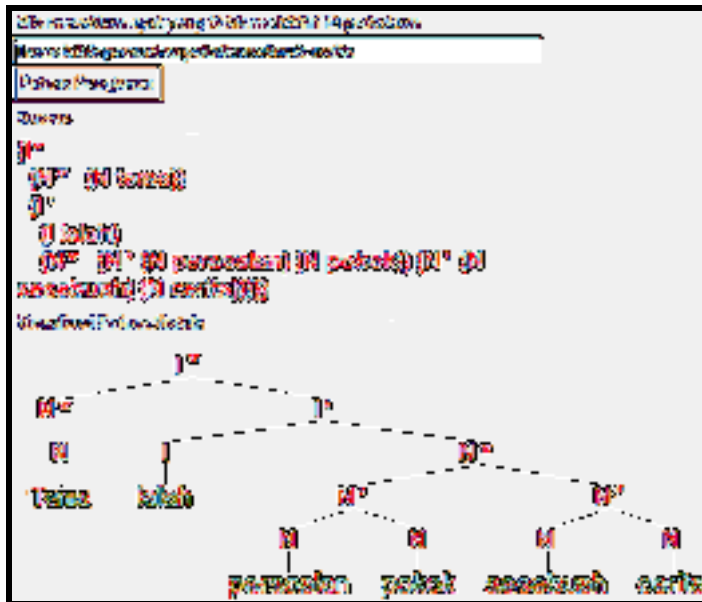
Rajah 5.17. Contoh VPS

Rajah 5.17 di atas hanya melibatkan satu frasa pertengahan (N') pada permulaan ayat. Kelas kata 'Pent' membawa maksud 'penentu' yang membantu menghuraikan frasa nama pertengahan (N'). Manakala kata adjektif yang menerangkan tentang subjek dalam ayat tersebut dikategorikan sebagai ayat berpola frasa adjektif (A"). Manakala kata bantu yang sepatutnya diwakili dalam ayat digugurkan. Ayat tersebut sebenarnya boleh berupa seperti "Pegawai polis itu sangat/amat/sungguh baik". Hal ini kerana frasa adjektif biasanya dipadankan dengan kata bantu atau kata penguat di awal ayat. Sebagai contoh, Rajah 5.18 menunjukkan contoh ayat yang mempunyai kata pemerii "ialah" di hadapan frasa nama (N").



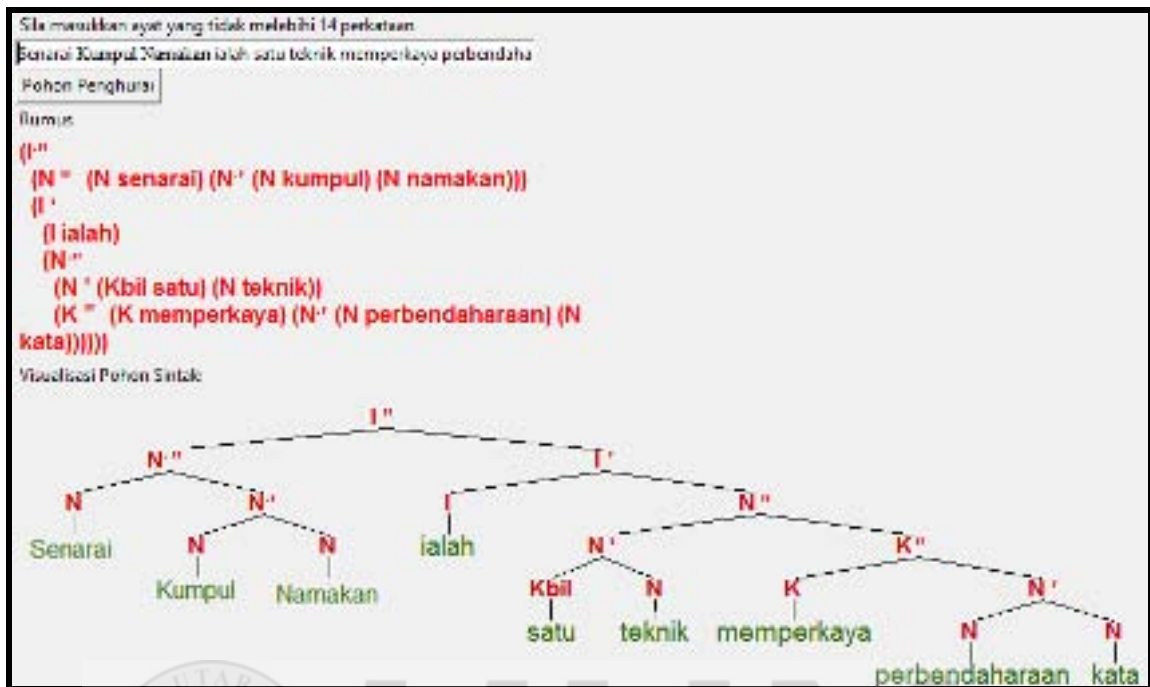
Rajah 5.18. Contoh VPS

Rajah 5.18 menunjukkan kata pemerii "ialah" yang tidak boleh digugurkan seperti kata bantu. Rajah 5.19 pula memberi contoh ayat berpola frasa nama (N") di mana terdapat dua frasa pertengahan (N') dengan gabungan antara kelas kata nama (N).



Rajah 5.19. Contoh VPS

Rajah 5.20 pula adalah contoh bagi ayat panjang melebihi tujuh perkataan. Ayat sebegini mempunyai gabungan rumus yang dibentuk dari jenis frasa dan kelas kata yang berbeza. Kedalaman dan kelebaran VPS boleh menjadi panjang dan lebar. Sebagai contoh, Rajah 5.20 mencapai tujuh tahap kedalaman untuk sembilan perkataan berbanding Rajah 5.18 yang mempunyai lima kedalaman untuk empat perkataan.



Rajah 5.20. Contoh VPS

Secara keseluruhan, dengan terhasilnya VPS, maka paparan atribut perkataan juga berjaya dilakukan kerana paparan ini hanya melibatkan penggunaan pautan yang diberi pada nod VPS. Semua ayat yang diuji dalam kajian ini adalah ayat yang betul dari segi sintaksis dan semantik. Seterusnya, setelah VPS berjaya diuji dan berjaya menghasilkan output yang positif iaitu mencapai lebih daripada 70% ketepatan VPS dianggap sebagai hasil yang boleh diterima (Tullis & Albert, 2013), maka proses seterusnya adalah untuk menguji ketepatan output untuk cadangan pembetulan ayat.

5.6.4 Hasil Uji Kaji Cadangan Pembetulan Ayat

Uji kaji untuk pemprosesan cadangan pembetulan ayat juga menggunakan ayat yang dikumpulkan dalam fasa ini. Bagi menjadikan ayat tersebut sebagai ayat yang salah, maka proses pengubahan kedudukan perkataan dibuat ke atas setiap ayat. Penilaian berdasarkan metrik *recall* dan *precision* juga digunakan. Hasil keseluruhan termasuk paparan antara muka yang dipaparkan ditunjukkan dalam Lampiran I.

Sebanyak 23 daripada 151 ayat tidak dapat memberi cadangan pembetulan yang betul. Ayat yang didapati salah cadangan yang diberikan adalah ayat yang mengandungi lebih daripada satu kata nama dalam frasa nama subjek, pengambilan hanya satu kata nama bagi ayat yang melebihi dari satu kata nama subjek akan mengeluarkan cadangan yang salah. Senarai ayat ini ditunjukkan dalam Jadual 5.9.

Jadual 5.9

Ayat dengan Cadangan yang Salah

Ayat dengan cadangan yang salah
1. Pegawai polis itu baik
2. Beberapa orang budak perempuan itu murid di sekolah saya
3. Kesannya saya rasa amat sukar bangun pagi
4. Kata penyambung ayat ialah kata hubung
5. Temenggung Abu_Bakar akhirnya diiktiraf sebagai Sultan Johor
6. Bangunan Sultan Ahmad_Samad merupakan bangunan tinggalan masa lampau
7. Negara asing pula tidak boleh campur tangan
8. Serangan wabak sars menyebabkan seluruh dunia gempar.
9. Pak Ismawan sering terasa Iswan ada di sisinya.
10. Penjaga garisan sudah mengangkat bendera kuning.
11. Bahan buangan ini dapat dikitar semula
12. Hari gawai menjadi perayaan utama di negeri saya.

Sambungan Jadual 5.9

13. Ayah Khairi berbincang dengan keluarganya
14. Surat tidak rasmi disebut juga sebagai surat kiriman biasa.
15. Syarikat tempat Hadi bekerja menyalahkan Hadi atas kerugian tersebut.
16. Mereka tidak mengambil peduli akan nilai sambutan hari kebangsaan
17. Dua orang guru sepenuh masa telah ditugaskan mengajar ahli kelab.
18. Rakan anda memberikan hujahnya tentang simpanan gas asli di negara kita
19. Seorang lelaki separuh umur masuk
20. Pelajar lelaki itu tersenyum
21. Pemain bola itu amat lincah
22. Pemandahan organ merupakan salah satu pencapaian sains perubatan moden
23. Engkerurai ialah sejenis alat muzik warisan

Sebanyak enam (6) daripada 128 ayat cadangan yang betul, dikategorikan betul oleh penilai Puan Siti Salmah walaupun ayat-ayat ini tidak mengikut ayat sasaran. Ayat ini dinilai betul dari segi penggunaannya. Ringkasan hasil uji kaji ini ditunjukkan dalam Jadual 5.10.

Jadual 5.10

Cadangan Pembedulan Ayat

Ayat uji kaji	Ayat salah	Cadangan pembedulan
Keindahan Putrajaya ini saya kagumi	Kagumi keindahan Putrajaya ini saya	Saya kagumi keindahan Putrajaya ini
Kata penyambung ayat ialah kata hubung	Ialah kata hubung kata penyambung ayat	Ayat ialah kata hubung kata penyambung
Masyarakat Sarawak menarik tarian ini	menarik tarian ini Masyarakat Sarawak	Sarawak menarik tarian ini Masyarakat
Kejayaan polis itu melegakan orang ramai	Melegakan kejayaan polis itu orang ramai	Orang ramai melegakan kejayaan polis itu
Keahlian kelab ini terhad	Terhad keahlian kelab ini	Kelab ini terhad keahlian
Reka bentuk masjid ini amat menarik	Amat menarik reka bentuk masjid ini	Masjid ini amat menarik reka bentuk

Berdasarkan kepada hasil uji kaji, nilai peratusan ketepatan dinilai berdasarkan metrik *recall* dan *precision* seperti yang dilakukan dalam menilai hasil uji kaji VPS. Sama seperti penilaian yang dibuat ke atas hasil uji kaji VPS, hasil uji kaji cadangan pembetulan ayat juga adalah berdasarkan kepada pola ayat yang diasingkan. Hasil uji kaji ini ditunjukkan dalam Jadual 5.11.

Jadual 5.11

Ringkasan Hasil Uji Kaji Cadangan Pembetulan Ayat

Pola ayat	Bilangan ayat terlibat	Hasil cadangan	Hasil cadangan yang betul
Frasa nama	18	18	14
Frasa kerja	98	98	83
Frasa adjektif	24	24	20
Frasa sendi nama	11	11	11
Jumlah ayat uji kaji	151	151	128

Setiap pola ayat dinilai mengikut kategori bagi kedua-dua metrik penilaian yang ditunjukkan seperti dalam Jadual 5.12.

Jadual 5.12

Purata dan Peratus Hasil Uji Kaji Cadangan Pembedulan Ayat

Pola ayat	Recall Jumlah output tepat/Output tepat yang dikira	Precision Output tepat/Semua ayat uji kaji
Frasa nama	14/14=100%	14/17=82.4%
Frasa kerja	83/83=100%	83/97=85.6%
Frasa adjektif	20/20=100%	20/24=83.3%
Frasa sendi nama	11/11=100%	11/11=100%
Purata	100%	87.8%

Hasil uji kaji cadangan pembedulan ayat seperti yang ditunjukkan dalam Jadual 5.12 di atas menunjukkan peratusan *precision* mencapai purata sebanyak 87.8%. Di antara keempat-empat pola ayat yang dibahagikan, peratusan penilaian untuk ayat frasa sendi nama lebih tinggi berbanding pola ayat yang lain dengan purata peratusan sebanyak 100% iaitu ketepatan penuh dalam menghasilkan cadangan pembedulan ayat. Manakala pola ayat frasa nama mencatat peratusan paling rendah iaitu 82.4% yang menunjukkan secara relatif prototaip sukar menganalisis ayat dalam pola ini berbanding ayat dalam pola yang lain.

5.6.5 Penilaian Pengguna

Hasil penilaian mengikut soal selidik *Usefulness, Satisfaction, and Ease of use (USE)* mendapat purata keseluruhan 87.9% dengan nilai min sebanyak 6.157 untuk penilaian penggunaan aplikasi VPS. Tujuan penilaian ini adalah untuk mengukur skor tahap

kepuasan subjektif pengguna sama ada hasil yang diperolehi memuaskan atau sebaliknya. Hasil keseluruhan mengikut soalan soal selidik ditunjukkan dalam Jadual 5.13.

Jadual 5.13

Hasil Keseluruhan Berdasarkan Soal Selidik USE

No	Soalan	Dimensi soalan	Skala Min	Skala Mak	Min (n=30; Skala 1-7)	Min faktor	Sisihan piawai
1	Ia boleh membantu saya menjadi lebih cekap	<i>Usefulness</i>	2	7	5.967	6.038	1.474
2	Ia membantu saya menjadi lebih berpotensi		1	7	5.733		1.507
3	Ia berguna kepada saya		3	7	6.400		1.037
4	Saya boleh mengawal penggunaan aplikasi dengan baik		2	7	5.733		1.574
5	Ia mengeluarkan output dengan mudah		3	7	6.167		1.289
6	Ia menjimatkan masa saya		1	7	6.133		1.408
7	Saya jumpa apa yang diperlukan		2	7	6.367		1.189
8	Ia melakukan apa yang diharapkan		2	7	5.800		1.472
9	Ia mudah digunakan		2	7	6.167		1.392
10	Ia senang digunakan		3	7	6.233		1.073
11	Ia mesra pengguna		4	7	6.100		6.042
12	Ia melibatkan turutan yang ringkas	<i>Ease of use</i>	2	7	5.933	1.437	
13	Ia kelihatan menarik		1	7	5.667	1.845	
14	Tidak susah menggunakannya		1	7	5.833	1.621	
15	Saya boleh menggunakannya tanpa manual bertulis		1	7	6.067	1.388	
16	Saya tidak melihat apa-apa yang tidak konsisten ketika saya menggunakannya		1	7	5.633	1.771	
17	Pengguna akan menyukai aplikasi ini		4	7	6.433	.898	
18	Saya boleh kembali daripada kesalahan dengan cepat dan mudah		3	7	6.367	1.129	
19	Saya boleh gunakan aplikasi ini		1	7	6.033	1.474	

Sambungan Jadual 5.13

No	Soalan	Dimensi soalan	Skala Min	Skala Mak	Min (n=30; Skala 1-7)	Min faktor	Sisihan piawai
	dengan jayanya						
20	Saya belajar menggunakannya dengan cepat		3	7	5.967	6.125	1.189
21	Saya mudah mengingat cara menggunakannya	<i>Ease of learning</i>	4	7	6.233		1.006
22	Saya mudah untuk belajar menggunakannya		4	7	6.233		1.073
23	Saya pandai menggunakannya tanpa bantuan		3	7	6.067		1.143
24	Saya berpuas hati dengan aplikasi ini		3	7	6.467		1.224
25	Saya akan cadangkan kepada kawan-kawan		3	7	6.467		.900
26	Ia menarik untuk digunakan		4	7	6.500		.820
27	Ia memberi output seperti yang diperlukan	<i>Subjective Satisfaction</i>	3	7	6.200	6.490	1.064
28	Ia adalah sebuah aplikasi yang baik		3	7	6.533		1.008
29	Saya berasa saya perlukan aplikasi ini		3	7	6.467		1.106
30	Saya berasa seronok menggunakan aplikasi ini		5	7	6.800		.484
Nilai min keseluruhan						6.157	

Secara keseluruhan, dimensi soalan kepuasan subjektif mencatat nilai min 6.490. Hal ini dapat dinyatakan bahawa kepuasan subjektif pengguna dengan menggunakan aplikasi BMTutor yang diberikan adalah tinggi. Nilai min bagi semua dimensi mencatat skala melebihi 6. Antara 30 soalan yang diajukan, soalan 16 mencatat min 5.633 iaitu nilai min paling rendah berbanding soalan yang lain. Soalan ini adalah berkenaan struktur maklumat yang konsisten. Nilai min tertinggi adalah bagi soalan terakhir iaitu sebagai rumusan penggunaan aplikasi yang diberikan. Soalan 30 ini mencatat nilai min sebanyak 6.800 yang menunjukkan pengguna berasa seronok menggunakan aplikasi yang

diberikan. Skala terendah bagi soalan ini juga menunjukkan skala 5 diberikan berbanding soalan yang lain.

Dimensi kebergunaan

Dimensi kebergunaan (*usefulness*) mencatat nilai min faktor 6.038 bagi soalan 1 hingga 8. Antara soalan ini, soalan 3 mencatat nilai min tertinggi iaitu 6.400 yang menunjukkan responden bersetuju bahawa aplikasi ini berguna kepada mereka. Soalan 2 dan 6 mencatat jurang yang besar antara skala 1 dan skala 7. Soalan ini adalah berkaitan dengan potensi dan masa yang menjimatkan daripada penggunaan aplikasi. Bagi soalan 2, skala 1 diberikan oleh pelajar tingkatan 1 dan soalan 6 pula skala 1 diberikan oleh pelajar tingkatan 4. Skala yang diberikan dapat dinyatakan bahawa aplikasi ini tidak membantu pelajar tingkatan 1 untuk menjadi lebih berpotensi. Walau bagaimanapun, Nilai min yang diperolehi untuk semua soalan dalam dimensi ini melebihi skala 5 yang menunjukkan responden bersetuju dengan kebergunaan aplikasi ini.

Dimensi kemudahan penggunaan

Dimensi kemudahan penggunaan (*ease of use*) mencatat min faktor 6.042 bagi soalan 9 hingga 19. Soalan 17 mencatat nilai min tertinggi iaitu 6.433 yang menunjukkan pengguna menyukai aplikasi ini. Terdapat lima soalan iaitu soalan (13,14,15,16,19) mencatat skala minimum 1 berbanding skala maksimum 7. Soalan ini berkaitan daya tarikan, kesukaran, bantuan penggunaan, penggunaan yang konsisten yang dirasakan

oleh pengguna dan kejayaan menggunakan aplikasi yang diberikan. Secara ringkas, Jadual 5.14 berikut menunjukkan soalan bagi skala minimum 1 yang diperolehi daripada penilaian.

Jadual 5.14

Soalan Bagi Skala Minimum 1 Soal Selidik USE

No	Soalan	Dimensi soalan	Skala Min	Pelajar
2	Ia membantu saya menjadi lebih berpotensi	<i>Usefulness</i>	1	Ting 1
6	Ia menjimatkan masa saya		1	Ting 4
13	Ia kelihatan menarik		1	Ting 1 dan 4
14	Tidak susah menggunakannya		1	Ting 1
15	Saya boleh menggunakannya tanpa manual bertulis	<i>Ease of use</i>	1	Ting 1
16	Saya tidak melihat apa-apa yang tidak konsisten ketika saya menggunakannya		1	Ting 2
19	Saya boleh gunakan aplikasi ini dengan jayanya		1	Ting 4

Berdasarkan Jadual 5.14, walaupun hanya seorang pelajar yang memberikan skala tersebut, namun dapat dinyatakan bahawa pelajar tingkatan 1 lebih sukar menerima aplikasi ini berbanding pelajar yang lain. Walau bagaimanapun, nilai min yang diperolehi melebihi dari skala 5 untuk kedua-dua dimensi.

Dimensi kemudahan pembelajaran dan kepuasan subjektif

Dimensi kemudahan pembelajaran (*ease of learning*) bagi soalan 20 hingga 23 mencatat min faktor 6.125. Dimensi ini mencatat skala minimum 3 yang menunjukkan responden boleh menerima aplikasi ini dan belajar menggunakannya dengan mudah. Dimensi

terakhir iaitu kepuasan subjektif (*subjective satisfaction*) juga mencatat skala minimum 3 dengan nilai min faktor 6.490. Hal ini menunjukkan kepuasan responden terhadap aplikasi ini memuaskan pengguna secara subjektif. Secara lebih jelas, hasil penilaian mengikut umur pelajar ditunjukkan dalam Jadual 5.15.

Jadual 5.15

Hasil Min Penilaian Soal Selidik USE Mengikut Tingkatan

Soalan (S)	Tingkatan 1	Tingkatan 2	Tingkatan 3	Tingkatan 4	Tingkatan 5
S1	5.000	7.000	6.500	4.667	6.667
S2	5.167	7.000	5.667	5.000	5.833
S3	6.667	7.000	6.333	5.500	6.500
S4	5.500	7.000	5.667	3.833	6.667
S5	6.000	7.000	6.333	5.000	6.500
S6	5.333	7.000	6.833	5.167	6.333
S7	6.167	7.000	6.167	5.667	6.833
S8	5.667	7.000	5.667	4.333	6.333
S9	5.167	7.000	6.000	5.667	7.000
S10	5.500	7.000	6.167	6.000	6.500
S11	5.167	7.000	6.167	5.333	6.833
S12	4.833	7.000	6.000	5.000	6.833
S13	4.833	7.000	6.000	4.167	6.333
S14	4.167	7.000	5.833	5.333	6.833
S15	4.167	7.000	6.500	6.167	6.500
S16	5.167	6.000	5.833	4.833	6.333
S17	5.833	7.000	6.000	6.333	7.000
S18	5.167	7.000	6.167	6.500	7.000
S19	5.500	7.000	5.833	4.833	7.000
S20	5.333	7.000	5.667	4.833	7.000
S21	5.833	7.000	5.833	6.000	6.500
S22	6.167	7.000	5.833	5.667	6.500
S23	5.500	7.000	6.167	5.333	6.333
S24	6.333	7.000	6.167	6.333	6.500
S25	6.333	7.000	6.333	6.000	6.667
S26	6.333	7.000	6.167	6.500	6.500
S27	6.667	7.000	6.167	4.833	6.333
S28	6.500	7.000	6.500	5.833	6.833
S29	6.833	7.000	6.000	5.667	6.833
S30	6.833	7.000	6.667	6.500	7.000
Nilai min	5.656	6.967	6.106	5.428	6.628

Nilai min bagi setiap tingkatan (Jadual 5.15) menunjukkan pelajar tingkatan dua mencatat tahap kepuasan subjektif lebih tinggi berbanding pelajar yang lain. Tingkatan satu dan empat mencatat purata pada skala 5. Penerimaan pelajar tingkatan empat lebih rendah berbanding pelajar yang lain. Pelajar tingkatan lima pula mencatat penerimaan yang tinggi dengan nilai min sebanyak 6.628. Oleh itu, dapat dinyatakan bahawa, purata skala melebihi 5 bagi semua dimensi menunjukkan penilaian yang dibuat berjaya mencapai objektif bagi mengukur tahap skor kepuasan subjektif pengguna apabila skor memuaskan diperolehi untuk setiap dimensi soalan. Namun, bagaimana pula dengan tahap penerimaan mereka? justeru itu penilaian penerimaan kognitif turut dilakukan.

Hasil penilaian kognitif secara keseluruhan mencatat nilai peratusan 84.6% dengan nilai min sebanyak 4.230. Semua 10 soalan yang diberikan mencatat nilai min melebihi skala 4 daripada 5. Tujuan penilaian kognitif ini adalah bagi mengukur tahap penerimaan pengguna. Hasil secara jelas ditunjukkan dalam Jadual 5.16.

Jadual 5.16

Hasil Penilaian Kognitif

No	Soalan	Skala Min	Skala Mak	Min (n=10; Skala 1-5)	Sisihan piawai
1	Semasa melihat paparan, saya cuba untuk membuat perkaitan antara maklumat paparan dengan konsep yang diajar oleh cikgu	2	5	4.200	.887
2	Semasa melihat paparan, saya cuba untuk kaitkan maklumat paparan dengan ilmu yang saya telah ketahui	2	5	4.067	.980
3	Semasa melihat paparan, saya cuba memahami idea dan	3	5	4.567	.626

Sambungan Jadual 5.16

No	Soalan	Skala Min	Skala Mak	Min (n=10; Skala 1-5)	Sisihan piawai
	konsep yang dipersembahkan				
4	Semasa melihat paparan, saya gabung maklumat yang telah diperolehi daripada pembelajaran di kelas dan pembacaan yang telah saya ketahui untuk memahami isi paparan secara keseluruhan	2	5	4.400	.894
5	Semasa paparan, saya cuba untuk membezakan maklumat utama (pohon sintaksis) dengan maklumat tambahan (atribut perkataan)	1	5	4.067	1.230
6	Semasa paparan, saya cuba untuk kaitkan maklumat paparan dengan subjek lain yang sealiran	2	5	4.133	.937
7	Saya cuba memahami jenis frasa dan cuba mengenal pasti perkaitan antara kelas kata terlibat	1	5	4.133	1.167
8	Saya memahami hasil paparan maklumat dari segi konsep dan idea persembahan	2	5	4.133	.937
9	Melihat akan paparan, saya terfikir untuk membuat paparan maklumat berkenaan dalam konsep persembahan yang lain	2	5	4.067	1.081
10	Ketika melihat paparan maklumat, ia membantu pemahaman saya tentang pembentukan ayat	3	5	4.533	.730
Nilai min keseluruhan				4.230	

Soalan 3 dan 10 berkaitan dengan pemahaman penggunaan mendapat respon yang tinggi. Skor ini menunjukkan pelajar cuba untuk memahami idea dan konsep yang dipaparkan dan dapat membantu meningkatkan pemahaman mereka dalam pembentukan ayat BM. Nilai min terendah pula diperolehi bagi item soalan 2, 5 dan 9. Respon yang diperolehi daripada soalan ini menunjukkan terdapat pelajar yang kurang berkebolehan untuk membuat perkaitan maklumat paparan dengan pengetahuan yang telah dipelajari di sekolah. Hal ini berkemungkinan bahawa, kaedah yang diperolehi hasil pembelajaran adalah berbeza. Walau bagaimanapun, semua soalan yang diajukan mendapat nilai min

melebihi daripada 4 yang menunjukkan mereka bersetuju dan memahami kaedah persembahan idea dan konsep yang dipaparkan.

Secara lebih jelas, hasil penilaian kognitif ini ditunjukkan mengikut umur pelajar seperti Jadual 5.17.

Jadual 5.17

Hasil Min Penilaian Kognitif Mengikut Tingkatan

Soalan (S)	Tingkatan 1	Tingkatan 2	Tingkatan 3	Tingkatan 4	Tingkatan 5
S1	4.167	4.833	4.167	3.167	4.667
S2	4.167	4.500	4.500	2.833	4.333
S3	4.667	4.833	4.333	4.333	4.667
S4	4.500	5.000	4.167	3.667	4.667
S5	3.833	4.833	4.500	2.833	4.333
S6	3.833	4.667	4.333	3.500	4.333
S7	3.833	5.000	4.167	3.667	4.000
S8	3.667	5.000	4.333	3.500	4.167
S9	3.667	4.667	4.000	3.333	4.667
S10	4.667	5.000	4.667	3.833	4.500
Nilai min	4.100	4.833	4.317	3.467	4.433

Hasil penilaian kognitif mengikut tingkatan (Jadual 5.17) menunjukkan pelajar tingkatan empat mencatat penerimaan keseluruhan pada skala 3 dengan nilai min 3.467. Hal ini menunjukkan pelajar berkenaan mempunyai tahap pemahaman dan penerimaan yang neutral. Manakala pelajar lain menunjukkan mereka bersetuju dan boleh memahami serta menerima paparan aplikasi yang ditunjukkan. Pelajar tingkatan dua lebih mudah memahami dan menerima hasil paparan dengan nilai min tertinggi iaitu 4.833 diikuti

pelajar tingkatan lima. Oleh itu, skor min yang melebihi skala 4 bagi semua soalan menunjukkan penerimaan kognitif pengguna adalah memuaskan. Mereka memahami dan menerima paparan yang dibuat dengan baik.

5.7 Perbincangan Dapatan

Peratusan ketepatan VPS adalah 4.6% lebih tinggi berbanding ketepatan output kajian Noor Hafhizah (2011) yang membuktikan rumus *X-bar* boleh menghasilkan VPS tepat yang lebih baik berbanding rumus binaan ayat BM yang lain seperti RSF. Ketepatan yang tinggi menyakinkan bahawa model dan algoritma yang diperkenalkan boleh digunakan untuk menghasilkan aplikasi sealiran yang lain. Model dan algoritma ini jika digunakan dalam kajian semakan semantik dipercayai akan menghasilkan ketepatan output yang hampir sama. Dalam semakan semantik, proses dan kaedah sama akan dilalui dengan penambahan turutan untuk semakan rumus.

Kajian pohon sintaksis terdahulu untuk ayat BM tidak melibatkan penilaian pengguna. Penilaian pengguna adalah penting untuk mengetahui penerimaan mereka dari segi persembahan visual dan penerimaan komponen VPS secara keseluruhan. Antara semua responden terlibat dalam penilaian pengguna, lima orang diambil yang mewakili setiap tingkatan untuk membuat rakaman aktiviti. Dalam tempoh penggunaan, pelajar lebih suka membuat paparan VPS dengan memasukkan ayat yang berlainan. Paparan tersebut dapat menarik minat mereka untuk mengetahui tentang pembentukan ayat dengan lebih

baik. Sokongan visual dalam paparan VPS tersebut sangat mempengaruhi minat mereka walaupun kandungan aplikasi belum tentu difahami. Perkara tersebut dibuktikan apabila semua pelajar terlibat melakukan semua turutan yang diperlukan dalam penggunaan aplikasi sehingga dapat menghasilkan VPS daripada ayat contoh.

Daripada hasil pengumpulan data yang dilakukan, pelajar tingkatan dua lebih mudah menerima dan memahami penggunaan aplikasi yang diberikan. Kesimpulan ini dibuat berdasarkan skor min tertinggi yang diperolehi daripada penilaian kepuasan subjektif dan kognitif. Pelajar tingkatan satu pula adalah pelajar yang didapati sukar menerima pembelajaran dan kebergunaan aplikasi yang diberikan. Skala minimum yang diperolehi bagi setiap dimensi soalan membuktikan perkara tersebut. Walaupun begitu, penerimaan pelajar tingkatan lima dalam penggunaan aplikasi ini adalah memuaskan. Kepuasan subjektif dan kognitif yang dicatatkan daripada pelajar tingkatan lima memberi skor min yang tinggi selepas pelajar tingkatan dua. Oleh yang demikian, dapat dinyatakan bahawa aplikasi yang dibangunkan menggunakan model dan algoritma VPS dengan output tambahan adalah memuaskan dan boleh membantu pelajar sekolah menengah terutama untuk tingkatan dua ke atas. Manakala bagi pelajar tingkatan satu pula, bantuan manual bertulis dan daya tarikan yang lebih menarik diperlukan untuk menggalakkan penerimaan mereka.

Justeru itu, hasil penilaian yang dilakukan telah menjawab persoalan kajian. Setiap persoalan ini dibincangkan dalam bahagian seterusnya.

Persoalan pertama "Adakah rumus binaan ayat telah diperkenalkan untuk VPS?"

Rumus binaan ayat sememangnya telah diperkenalkan dalam kajian pohon sintaksis terutama BM. Rumus ini digunakan dalam kajian Noor Hafhizah (2011), Ahmad Izuddin *et al.* (2007), Rosmah (1995) dan Suzaimah (2002). Namun, rumus yang digunakan adalah RSF yang dianggap tidak universal dalam kajian pemprosesan ayat. Kaedah pengumpulan rumus yang dilakukan dalam kajian tersebut tidak dinyatakan secara jelas, namun melalui hasil penulisan yang dilakukan, dapat diperolehi kaedah yang dibuat.

Kaedah pengumpulan rumus binaan ayat dalam kajian Noor Hafhizah (2011) diakui telah berjaya digunakan dalam kajian ini. Hasil analisis ke atas rumus binaan ayat BM telah menemukan rumus *X-bar* sebagai rumus yang terkini dalam kajian pemprosesan ayat BM. Rumus ini telah digunakan berdasarkan skop kajian dan berjaya menghasilkan VPS. Ini membuktikan bahawa rumus binaan ayat BM boleh digunakan untuk VPS dengan output tambahan. Oleh itu, Hipotesis pertama kajian mengenai kebolehan rumus binaan ayat dalam kegunaan VPS boleh dibuktikan bahawa rumus *X-bar* berjaya menghasilkan VPS dengan output tambahan dengan peratusan *recall* paparan output VPS sebanyak 100%.

Rumus binaan ayat yang dicadangkan dalam kajian ini tidak hanya boleh digunakan dalam kajian pohon sintaksis. Rumus ini boleh ditambah elemen lain dan diubah berdasarkan sesebuah bahasa. Rumus ini jika digunakan dalam aplikasi pengecaman ayat dalam teks yang panjang, ia boleh memainkan peranan sebagai penanda struktur frasa. Sebagai contoh, ayat input berbentuk teks digunakan, atur cara semakan boleh dilakukan untuk membuat penanda frasa dan ayat dipaparkan kepada pengguna mengikut jenis. Biasanya, kaedah ini diperlukan dalam kajian mesin terjemahan.

Persoalan kedua "Bagaimana model, algoritma dan prototaip output tambahan pohon sintaksis boleh dibina".

Kaedah pembangunan aplikasi kajian terdahulu dianalisis bagi mendapatkan model, algoritma dan kaedah paparan output. Kaedah ini dimulai dengan menyenaraikan semua komponen model. Komponen ini dijadikan input untuk mereka bentuk turutan algoritma dengan output tambahan berdasarkan kepada skop kajian. Turutan yang diperolehi dijadikan panduan dalam membina prototaip BMTutor.

Penghasilan model, algoritma dan prototaip kajian dengan output tambahan membuktikan komponen ini boleh direka bentuk. Penambahan output berjaya dimodelkan dengan memasukkan elemen semakan, cadangan, VPS dan atribut

perkataan. Pembuktian konsep ke atas output tambahan ini berjaya menghasilkan output VPS yang tinggi berbanding kajian terdahulu. Oleh itu, hipotesis kajian yang kedua mengenai keupayaan model untuk digunakan dalam reka bentuk algoritma dan pembangunan prototaip boleh dibuktikan bahawa model VPS dengan output tambahan memenuhi syarat sebuah aplikasi VPS. Komponen model dipersetujui dan disahkan oleh pakar BM dan pakar bidang. Komponen ini selari dengan turutan algoritma kajian di mana model dan algoritma ini diterjemahkan dalam bentuk pembangunan prototaip. Keupayaan model dan algoritma ini terbukti memenuhi syarat sebuah VPS dengan output tambahan apabila berjaya mencapai hasil ketepatan output yang tinggi.

Sebuah model VPS sememangnya perlu dimuatkan komponen semakan dengan cadangan ayat. Turutan algoritma yang berdasarkan kepada model ini juga perlu dimasukkan elemen tersebut. VPS yang dihasilkan bukan hanya untuk membuat paparan output semata, namun dengan adanya elemen tersebut, paparan output yang dihasilkan akan memberi idea kepada pengguna untuk lebih memahami penggunaan ayat yang gramatis. Elemen ini bukan hanya boleh dimasukkan dalam VPS, malah boleh digunakan dalam kajian terjemahan ayat. Sebagai contoh, semakan padanan rumus boleh dibuat dalam konteks berbeza dan ayat yang dihasilkan dipaparkan sebagai cadangan kepada pengguna. Oleh itu, dapat dinyatakan bahawa elemen semakan dengan cadangan yang diberikan dalam

kajian ini menyumbang kepada penghasilan output tepat yang tinggi berbanding kajian terdahulu. Pengguna bukan hanya didedahkan tentang pembentukan ayat, malah penggunaan ayat mengikut rumus boleh memberi ilmu pengetahuan tentang struktur pembentukan ayat yang baik.

Output tambahan berupa VPS pula sangat membantu pemahaman pembentukan ayat. Berbeza dengan kajian terdahulu apabila elemen visualisasi diambil kira. Elemen ini memberi bantuan kepada pengguna dengan menyediakan pautan (*hyperlink*) untuk menerokai pembentukan ayat dengan lebih baik. Elemen ini perlu dimasukkan dalam pohon sintaksis agar output yang dipaparkan tidak hanya memberi gambaran semata, namun dapat memberi bantuan pemahaman dari segi penggunaan kelas kata, kata terbitan, terjemahan, imej dan ayat contoh. Elemen ini juga boleh digunakan dalam aplikasi lain, sebagai contoh, aplikasi semakan kelas kata, semakan semantik, ringkasan teks dan mesin terjemahan. Pautan boleh diberikan dalam aplikasi tersebut untuk memaparkan atribut berkaitan dan kaedah ini dipercayai akan lebih memudahkan pemahaman ke atas sesebuah perkataan dan ayat. Hal ini kerana penentusahan komponen model menunjukkan komponen ini sangat tepat penggunaannya dalam pembelajaran pembentukan ayat BM kerana ia berdasarkan kepada rumus *X-bar*.

Persoalan kajian ketiga "Bagaimana cara untuk memastikan model dan algoritma yang dihasilkan adalah tepat?"

Kaedah pembuktian konsep yang dilakukan dalam kajian terdahulu terutama kajian Ahmad Izuddin *et al.* (2007) dan Noor Hafhizah (2011) dijadikan panduan dalam membentuk kaedah pembuktian konsep kajian ini. Kajian tersebut menggunakan metrik penilaian Parseval untuk mengukur nilai skor paparan output. Oleh itu, kaedah tersebut digunakan dalam kajian ini. Pembuktian konsep yang dilakukan menunjukkan persoalan ketiga berjaya dijawab. Hasil ketepatan output menunjukkan kaedah *gold-reference* iaitu kaedah semakan hasil penilaian secara bandingan output manual dan sistem mempunyai hubungan kepentingan yang kuat. Hasil ketepatan yang diperolehi hampir tepat untuk semua ayat mengikut output manual yang dibandingkan. Hasil output ini adalah tinggi berbanding karya terdahulu. Hasil ini dapat membuktikan hipotesis kajian yang ketiga bahawa skor kepuasan subjektif dan kognitif pengguna dapat diukur.

Secara keseluruhan, untuk membangunkan sebuah VPS, proses yang berlaku dianggap tidak lengkap untuk menganalisis ayat jika tidak melibatkan elemen semakan, cadangan dan atribut perkataan. VPS dengan output tambahan ini telah dibuktikan boleh menghasilkan output yang tepat sebagaimana yang dihasilkan secara manual. Ia boleh menyediakan panduan kepada pembangun perisian kajian

pemrosesan ayat dalam bidang pengkomputeran linguistik supaya dapat memberi kesan penggunaan kepada pelajar sebagai pengguna.

5.8 Rumusan Bab Lima

Bab ini telah membincangkan tentang proses pembangunan prototaip bermula dari lakaran antara muka secara manual sehingga dapat dihubungkan dengan pangkalan data. Hasilnya paparan antara muka yang lengkap turut dijelaskan mengikut turutan algoritma VPS dengan output tambahan. Selain itu, setiap fail pangkalan data yang terlibat dalam kajian ini seperti fail rumus, fail perkataan, dan fail ayat contoh ditunjukkan contoh maklumat disimpan dan dijelaskan perkaitan dan keperluannya.

Fasa latihan prototaip dilakukan untuk mendapatkan rumus yang lasak. Proses ulangan dalam fasa latihan ini dilakukan sehingga mendapat output VPS tepat yang tinggi. Kekangan yang dikenal pasti, ditambah baik dalam penggunaan rumus sehingga dapat menghasilkan output VPS yang tepat sebanyak 82 daripada 106 ayat.

Seterusnya, fasa penilaian prototaip diteruskan dengan menggunakan ayat uji kaji yang berlainan. Dapatan daripada penilaian prototaip ditunjukkan dalam bentuk metrik penilaian Parseval dan penilaian pengguna. Nilai peratusan ketepatan (*precision*) yang diperolehi adalah sebanyak 97.8% untuk uji kaji dari segi VPS iaitu 147 ayat berjaya membuat VPS dengan tepat daripada 151 ayat uji kaji. Manakala hasil dari segi

cadangan pembetulan ayat pula mencapai peratusan ketepatan sebanyak 87.8% iaitu 128 ayat berjaya memberi cadangan pembetulan yang betul daripada 151 ayat keseluruhan. Penilaian pengguna pula mencatat peratusan 87.9% untuk penilaian kepuasan subjektif pengguna dengan nilai min sebanyak 6.157. Penilaian kognitif pula mencatat peratusan 84.6% dengan nilai min sebanyak 4.230.

Dapat disimpulkan bahawa penilaian yang dilakukan telah membuktikan konsep dalam model dan algoritma kajian ini sememangnya telah menghasilkan output yang positif. Hal ini kerana output melebihi 70% dianggap sebagai boleh diterima (Tullis & Albert, 2013). Hasil penilaian ini telah menjawab ketiga-tiga persoalan kajian yang dibincangkan dalam bahagian dapatan kajian. Seterusnya, untuk melihat perkaitan pencapaian objektif dengan setiap aktiviti yang dilakukan dalam kajian ini, Bab Enam seterusnya memberi rumusan keseluruhan termasuklah kekangan dan penambahbaikan di masa akan datang.

BAB ENAM

RUMUSAN

6.0 Pengenalan

Bab ini membuat rumusan secara keseluruhan yang meliputi penjelasan mengenai sumbangan kajian, pencapaian objektif, kekangan dan penambahbaikan yang boleh dilakukan di masa akan datang. Sumbangan yang dijelaskan berupa rumus binaan ayat, model dan algoritma serta atribut perkataan. Kekangan daripada sumbangan ini dikemukakan agar penambahbaikan boleh dilakukan pada masa akan datang.

6.1 Rumusan Sumbangan Kajian

Kajian terdahulu terhad kepada output pohon sintaksis sebagai output akhir. Tambahan elemen output dalam kajian ini meningkatkan lagi keupayaan pohon sintaksis dalam menghasilkan ketepatan output. Hasil dapatan kajian membuktikan ketepatan output VPS melebihi tahap ketepatan output pengkaji terdahulu. Oleh itu, tambahan elemen output ini memberi sumbangan seperti berikut.

I. Model VPS dengan output tambahan

Model terdahulu melibatkan komponen asas penghurai ayat (rujuk bahagian 2.4.1). Output tambahan yang dicadangkan meliputi cadangan ayat, VPS dan atribut perkataan adalah komponen baru yang digabung dalam VPS. Gabungan komponen ini menjadikan VPS bukan hanya dapat memaparkan pohon sahaja,

tetapi output lain boleh dimuatkan hasil daripada penggunaan satu perkataan (rujuk bahagian 5.3).

Model ini mencadangkan semakan dan cadangan ayat mengikut padanan rumus bagi struktur kelas kata dan jenis frasa. Cadangan ayat yang dilakukan bukan hanya untuk membantu binaan ayat yang betul tetapi untuk menggalakkan penggunaan ayat yang gramatis. Sumbangan ini penting untuk menyelesaikan masalah dalam bidang berkaitan dengan struktur pemprosesan ayat seperti bidang pengkomputeran linguistik.

II. Algoritma VPS dengan semakan serta cadangan

Turunan asas pemprosesan bahasa tabii digunakan dalam karya terdahulu meliputi tiga turutan. Tambahan turutan algoritma dicadangkan berdasarkan model dan skop kajian yang meliputi 10 turutan (rujuk bahagian 4.2.1). Turutan ini ditunjukkan setiap output tambahan yang terdapat dalam model VPS. Penambahan turutan yang dicadangkan boleh digeneralisasikan dalam penggunaan VPS dan aplikasi sealiran seperti penyemak semantik dan terjemahan ayat.

III. Rumus binaan ayat

Rumus *X-bar* yang digunakan dalam VPS ayat BM masih baru dikaji. Kajian ini menganalisis penggunaan rumus ini bagi kegunaan ayat penyata BM (rujuk Bab

5). Ketepatan output penilaian yang dilakukan menunjukkan rumus ini berjaya digunakan dalam VPS (rujuk bahagian 5.5). Penghasilan output juga dapat dilakukan untuk semua ayat menggunakan rumus ini berbanding karya terdahulu yang tidak dapat menghasilkan output untuk semua ayat uji kaji (rujuk bahagian 5.6.3).

IV. Pengecaman atribut perkataan

Setiap perkataan dalam VPS boleh menghasilkan atribut perkataan bagi lima atribut berbeza. Atribut ini boleh menghasilkan VPS melalui ayat contoh. Tambahan output ini memberi gambaran imej sebagai bantuan pemahaman. Tambahan output ini juga tidak terdapat dalam karya terdahulu (rujuk bahagian 2.4.6), dan telah dibuktikan bahawa output ini telah berjaya dimodelkan dalam VPS. Hal ini kerana output ini berjaya disertakan melalui pautan pada nod VPS dan juga dapat membuat paparan VPS ayat contoh (rujuk bahagian 5.3.4).

6.2 Rumusan Pencapaian Objektif

Bagi mencapai objektif utama kajian untuk menghasilkan model dan algoritma VPS dengan output tambahan, maka sub-objektif terlebih dahulu dicapai. Pencapaian sub-objektif ini dijelaskan dalam Bab Satu hingga Bab Lima.

Objektif 1: Mengkategorikan dan menganalisis struktur ayat BM untuk mendapatkan rumus yang tepat.

Penjelasan tentang objektif pertama dapat dilihat dalam Bab Tiga iaitu bahagian pengumpulan ayat BM mengikut skop kajian. Dalam proses pengumpulan ayat, terdapat beberapa proses yang dilalui antaranya penentuan skop ayat, pengumpulan ayat daripada buku teks BM tingkatan satu hingga tingkatan lima, lakaran pohon sintaksis dan pengesahan rumus untuk tujuan pembangunan, latihan dan uji kaji prototaip.

Berdasarkan analisis sorotan karya, rumus binaan ayat BM telah dihasilkan oleh Noor Hafhizah (2011). Kaedah pengumpulan rumus yang dilakukan dijadikan panduan untuk mencapai objektif pertama.

Objektif 2: Membina model dan algoritma VPS dengan output tambahan serta prototaip kajian sebagai alat pembuktian konsep.

Penjelasan tentang objektif kedua boleh dirujuk dalam Bab Empat. Model atribut perkataan terdiri daripada kelas kata, kata terbitan, terjemahan, imej dan ayat contoh bagi setiap perkataan yang dipilih semasa paparan VPS ayat input. Model ini digabung dalam VPS dengan menghasilkan model VPS dengan output tambahan.

Penentusahan model dilakukan oleh lima orang pakar terdiri daripada dua orang pakar bidang pengkomputeran linguistik dan tiga orang Munsyi Dewan. Penjelasan tentang

hasil penilaian dan cadangan disertakan dalam Bab Empat. Hasil penilaian menunjukkan komponen model dipersetujui oleh pakar penilai untuk membuat VPS.

Daripada sepuluh langkah yang terdapat dalam algoritma VPS dengan output tambahan, kaedah yang terlibat dalam membina algoritma penyemak dan cadangan pembetulan ayat BM melibatkan tujuh kaedah yang pertama iaitu kaedah token perkataan, semak bilangan perkataan, semak syarat ayat, padanan kelas kata, semak ejaan, semak rumus dan proses cadangan. Tiga kaedah terakhir iaitu VPS, atribut perkataan dan VPS ayat contoh adalah algoritma yang terlibat dalam membangunkan algoritma VPS struktur ayat BM. Seperti gabungan yang dilakukan dalam pembangunan model, algoritma VPS dengan output tambahan juga dihasilkan dengan menggabungkan algoritma yang direka bentuk. Turutan kaedah pelaksanaan atribut perkataan dalam VPS juga digabung dalam algoritma ini.

Objektif 3: Menilai output VPS dan cadangan pembetulan ayat berdasarkan metrik penilaian dalam pemprosesan bahasa tabii untuk menguji ketepatan output dan membuat penilaian pengguna bagi mencapai kepuasan subjektif dan penerimaan kognitif pengguna.

Hasil dan proses yang terlibat dalam menjalankan fasa penilaian boleh dirujuk dalam Bab Lima. Sebelum uji kaji sebenar prototaip dilakukan, fasa latihan prototaip

dilaksanakan. Tujuan fasa latihan ini adalah untuk mengenal pasti kekangan dan penambahbaikan yang boleh dilakukan terutamanya dalam penggunaan rumus.

Setelah semua objektif tercapai, maka dapat disimpulkan bahawa kajian ini berjaya memberi sumbangan baharu dalam bidang teknologi maklumat khususnya bidang pengkomputeran linguistik. Walau bagaimanapun, terdapat kekangan yang telah dikenal pasti yang boleh memberi ruang kepada penambahbaikan pada masa hadapan terutama jika kajian ini hendak diteruskan atau hendak dilakukan dalam konteks domain yang lain. Setiap kekangan ini dijelaskan dalam bahagian seterusnya.

6.3 Kekangan Kajian

Kajian ini juga terdapat kekurangan atau kekangan untuk mencapai kesempurnaan seperti kajian lain. Antara kekangan yang dikenal pasti ialah masalah kekaburan, bilangan perkataan yang terhad, penggunaan ayat yang terhad dan penggunaan contoh ayat yang terhad.

Kajian ini tidak menitikberatkan aspek semantik dan hanya melibatkan aspek sintaksis. Oleh sebab itu, output yang dihasilkan bagi perkataan yang kabur, akan mempunyai lebih daripada satu paparan VPS mengikut jumlah kelas kata yang dikenal pasti oleh prototaip. Kekaburan makna perkataan ini akan menyumbang juga kepada kekaburan

struktur ayat. Walaupun satu perkataan mempunyai dua kelas kata, tetapi paparan VPS boleh melebihi jumlah tersebut kerana susunan rumus yang dimiliki.

Namun, hasil kajian ini juga adalah untuk membantu penutur BM memahami struktur pembentukan ayat BM dengan lebih baik. Walaupun masalah paparan VPS yang kabur akan melebihi daripada satu paparan, jika dilihat dari segi kebaikannya, paparan tersebut sebenarnya mempunyai kepentingan tersendiri. Pengguna bukan sahaja dapat mempelajari bahawa satu perkataan tidak hanya memiliki satu kelas kata dan dapat membantu pengguna untuk mempelajari perkataan kabur tersebut untuk digunakan dalam konteks ayat yang lain.

Selain itu, pengguna juga dapat mengetahui penggunaan struktur ayat yang berbeza dengan menggunakan satu perkataan. Proses pelajaran tidak membataskan pengetahuan pengguna sebaliknya menunjukkan setiap aspek agar lebih difahami. Oleh sebab itu, aspek semantik tidak dikaji. Hal ini kerana jika aspek semantik dikaji, pengguna akan dipaparkan dengan hanya satu VPS yang mempunyai tahap ketepatan pohon yang lebih baik daripada pohon kabur yang lain. Jika dilihat dari segi kebaikan, sememangnya kaedah secara semantik dapat membantu pengguna agar tidak keliru dan akan mendapat output yang terbaik. Namun, pembelajaran tidak akan dilakukan secara menyeluruh ekoran pengguna tidak didedahkan dengan perkataan atau struktur ayat yang berbeza. Perkara tersebut akan membuatkan pengguna menjadi lebih mudah keliru apabila

berinteraksi dengan ayat dan perkataan yang kabur kerana pengguna tidak mempunyai pengetahuan tentang kelas kata kedua atau yang lain yang dimiliki oleh setiap perkataan kabur. Konteks ayat yang dapat dipelajari juga tidak akan meluas.

Kajian ini masih di peringkat permulaan dan skop bilangan perkataan juga adalah terhad. Jika perkataan yang digunakan melebihi skop ayat, maka mesej ralat diberikan. Ini menjadikan prototaip hanya boleh digunakan untuk ayat tertentu mengikut skop kajian dan menjadi satu kekangan untuk pengguna memasukkan ayat yang panjang.

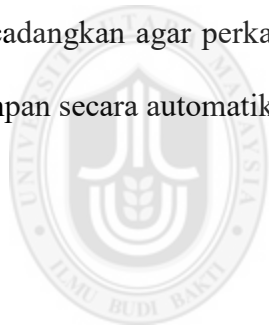
Ayat contoh yang diberikan dalam paparan atribut perkataan adalah ayat yang telah ditetapkan dalam pangkalan data. Ini memberi kekangan kepada pengguna untuk mempelajari kepelbagaian ayat yang lain. Selain itu, bilangan ayat contoh yang diberikan pada paparan antara muka utama adalah ayat yang terdapat dalam buku teks tingkatan satu hingga tingkatan lima yang terhad.

Kekangan yang dikenal pasti dapat ditambah baik untuk kegunaan kajian lain yang sealiran. Penambahbaikan yang boleh dilakukan dijelaskan dalam bahagian seterusnya.

6.4 Penambahbaikan Masa Hadapan

Kajian ini boleh diteruskan pada masa akan datang untuk menghasilkan prototaip yang lebih baik. Algoritma boleh ditambah untuk mengurangkan kekaburan perkataan atau

struktur ayat dengan menggunakan nilai kebarangkalian dalam penggunaan rumus. Skop kajian juga boleh diperluaskan agar tidak hanya ayat penyata yang dianalisis. Jenis ayat lain boleh dikaji sebagai penambahan. Selain itu, had bilangan perkataan boleh diperluaskan menjadi teks yang panjang agar VPS boleh dibuat dengan lebih mendalam agar pengguna lebih mendapat pengalaman dan pengetahuan. Ayat contoh dalam atribut perkataan juga boleh diambil terus daripada laman sesawang atau korpus ayat yang berbeza agar capaian boleh dibuat secara automatik dan bukannya telah ditetapkan dalam pangkalan data seperti yang dilakukan dalam kajian ini. Senarai contoh ayat juga boleh ditambah mengikut konteks ayat yang pelbagai. Selain itu, kajian ini juga mencadangkan agar perkataan yang tidak terdapat dalam pangkalan data ditambah dan disimpan secara automatik oleh prototaip.



UUM
Universiti Utara Malaysia

RUJUKAN

- Abd. Aziz, A. T. (2000). *Pedagogi Bahasa Melayu, prinsip, kaedah, dan teknik*. Utusan Publications & Distributors Sdn. Bhd.: Kuala Lumpur.
- Abdul Rahman Huraissen, M. (2012). *Teori Pemprosesan Maklumat Berbantuan Perisian Multimedia*. Retrieved October 12, 2014, from http://www.polimas.edu.my/web4/images/inovasi/diges%201%20all%20pg%0num_141012.pdf.
- Abdul Rashid, D. M. (2004). *Perubahan struktur kata tunggal Bahasa Melayu mengikut aliran*. PhD Thesis. Universiti Putra Malaysia.
- Abdul Razif, Z., & Rosfazila, A. R. (2016). *Dato Mahamud bin Abdul, Tokoh Sosiolinguistik Melayu Nusantara: Suatu kajian Konseptual, Profil dan Sumbangan*. Retrieved December 13, 2017, from <file:///C:/Users/Yusnita/Downloads/ROSAZILA2015kajiansosiolinguistikDatoMahamudBinAbdul1.pdf>
- Abdullah, H. (2008). *Tatabahasa Pedagogi untuk sekolah menengah*. PTS Professional Publishing Sdn. Bhd.: Selangor.
- Abdullah, H., Seri Lanang, J. R., Razali, A., & Zulkifli, O. (2006). *Sintaksis, siri pengajaran dan pembelajaran Bahasa Melayu*. PTS Professional Publishing Sdn. Bhd.: Kuala Lumpur, Malaysia.
- Abdullah. (2010). *Pelan strategik untuk DBP*. Retrieved January 24, 2011, from <http://semanggol.com/index.php?view=article&catid=37:bicara-pendeta&id=467:pelelan-strategik-untuk-dbp&format=pdf>.
- Abrak, O. (2005). *Kamus Komprehensif Bahasa Melayu*. Fajar Bakti Sdn. Bhd: Shah Alam.
- Ahmad Izuddin, Z. A., Yong, S. P., Rozana, K., & Hazreen, A. (2007). Utilizing top-down parsing technique in the development of a Malay language sentence parser. *Proceeding of the 2nd International Conference on Informatics*. Kuala Lumpur.
- Ahmad Rizal, M., & Yahya, B. (2006). *Keupayaan Visualisasi Dan Gaya Kognitif Pelajar Melalui Perisian Multimedia*. Seminar Kebangsaan Pendidikan Teknik dan Vokasional 2006, Senai, Johor.
- Al-Adhaileh, M. H., & Kong, T. E. (1998). A flexible example-based parser based on the SSTC. *Proceeding COLING '98 Proceedings of the 17th international conference on Computational linguistics*, 687-693.
- Almeida-Martinez, F. J., Urquiza-Fuentes, J., & Velquez-Iturbide, A. (2009). Visualization of Syntax Trees for Language Processing Courses. *Journal of Universal Computer Science*, 15(7), 1546-1561.
- Alsaffar, A., & Nazlia, O. (2015). Integrating a Lexicon Based Approach and K Nearest Neighbour for Malay Sentiment Analysis. *Journal of Computer Science*. 11(4). 639.644.

- Azizah, A. (2012). *Hubungan tahap visualisasi, strategi kognitif, metakognitif, dan kebolehan spatial dengan tahap pencapaian penyelesaian masalah matematik berayat dalam kalangan pelajar tingkatan empat di Melaka, Malaysia*. Master Thesis. Universiti Putra Malaysia: Bangi, Malaysia.
- Azizi, Y., Asmah, S., Zurihanmi, Z., & Fawziah, Y. (2005). *Aplikasi kognitif dalam pendidikan*. PTS Professional Publishing Sdn. Bhd.: Pahang.
- Bagavathy, A. C. (2005). Mengatasi Kelemahan Murid Menguasai Aspek Tatabahasa Dalam Bahasa Melayu Melalui Cara Permainan Bahasa. *Prosiding seminar penyelidikan pendidikan IPBA 2005*, 50-58.
- Barr, V. (2003). A proposed model for effective verification of natural language generation systems. *Proceedings of the Sixteenth International Florida Artificial Intelligence Research Society Conference 2003 (FLAIRS 2003)*. 208-212.
- Barnett, V. (2002). *Sample survey: Principles & methods*. Third Ed. Oxford University Press Inc.: New York.
- Barr, V. B., & Klavan, J.L. (2001). *Verification and Validation of Language Processing Systems: Is It Evaluation?* ACL 2001 Workshop on Evaluation Methodologies for Language and Dialogue Systems. Toulouse. 34-40.
- Bastings, J., & Sima'an, K. (2014). All Fragments Count in Parser Evaluation. *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014)*. 78-82.
- Battista, G. D., Eades, P., Tamassia, R., & Tollis, I. G. (1999). *Graph Drawing: Algorithms for the visualization of graphs*. Prentice Hall: New Jersey.
- Behrenberg, C. (2009). *SynView v0.3 user's manual*. Retrieved December 22, 2010, daripada http://www.christianbeherenberg.de/files/SynView/SynView_source.rar.
- Bjork, S., Holmquist, L., & Redström, J. (n.d). *A framework for focus + context visualization*. Retrieved December 22, 2010, from <http://www.sics.se/fal/publications/play/2000/dissertations/leh/framework.pdf>.
- Bolshakov, I., & Gelbukh, A. (2004). *Computational Linguistics: Models, Resources, Application*. First Edition. Instituto Politécnico Nacional: Mexico.
- Charniak, E. (2000). *A maximum-entropy-inspired parser*. In Proceedings of the NAACL 2000. Seattle, Washington, April 29-May 3. 132-139.
- Chomsky, N. (1986). *Knowledge of Language: Its Nature, Origin and Use*. New York: Praeger.
- Chomsky, N. (1970). *Remarks on Nominalization*. In Jacobs, R. A., & Rosenbaum, P. S. (eds.). *Readings in English Transformational Grammar*. 184-221. Boston: Ginn
- Chomsky, N. (1957). *Syntactic Structures*. The Hague, Paris: Mouton.
- Chu Min Xian, B., Lubani, M., Kwei Ping, L., Bouzekri, K., Mahmud, R., & Lukose, Dickson. (2016). Benchmarking Mi-POS: Malay part-of-speech tagger. *International Journal of Knowledge Engineering*, 2(3).
- Chuah, C. K., & Zaharin, Y. (2002). Computational linguistics at Universiti Sains Malaysia. *Proceedings of Third International Conference On Language Resources and Evaluation (LREC-2002)*. Spain. 29 May-31 May.

- Cleven, A., Gubler, P., & Huner, K. M. (2009). Design alternatives for the evaluation of design science research artifacts. *DESRIST '09 Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology*. Doi: 10.1145/1555619.1555645.
- Collins, M. (2000). *Discriminative reranking for natural language parsing*. In Proceedings of ICML 2000. Stanford University, Palo Alto, CA, June 29-July 2.
- Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2001). *Introduction to Algorithms* (2nd Edition). Mit Press: Cambridge, London.
- Dougherty, R. C. (n.d). *Lehner's Prolog Tree Drawing*. Retrieved December 17, 2017, from <http://www.nyu.edu/pages/linguistics/workbook/lehner>
- Eisenbach, M., & Eisenbach, A. (2003). *phpSyntaxTree - drawing syntax trees made easy*. Retrieved December 20, 2010, from <http://www.ironcreek.net/phpsyntaxtree/>.
- Erfan, M., & Lili, A.N. (2014). Visualization of Subjective Extracted Text Using the Parse Tree. *International Journal of Computer and Information Technology*. 3(2). 333-340.
- Grinstein, G.G., & Ward, M.O. (2002). *Introduction to data visualization*. In Information Visualization in Data Mining and Knowledge Discovery. Editors: Fayyad, U., Grinstein, G. G., & Wierse, A. 21-22. Morgan Kaufmann Publishers: USA.
- Hamidah, C. M. (2010). *Kesan Pembelajaran Terhadap Prestasi Kemahiran Berfikir Kritis dalam Penulisan Rumusan bagi Subjek Bahasa Melayu di kalangan Pelajar Tingkatan Satu yang Berbeza Pencapaian*. Master Thesis. Universiti Sains Malaysia, Pulau Pinang, Malaysia.
- Hevner, A. R., March, S. T., & Park, J. (2004). Design Research in Information Systems Research. *MIS Quarterly*, 28(1), 75-105.
- Hicks, M. (2009). *Perceptual and design principles for effective interactive visualisations*. In Trend in Interactive Visualization, state-of-the-art survey. Editors: Zudilova-Seinstra, E., Adriaansen, T., & Liere, R. V. Springer: London.
- Hirschman, L., & Mani, I. (2004). *Evaluation*. In Mitkov, R. (eds). The oxford handbook of computational linguistics. 414-425. Oxford University Press: New York.
- Hunter, J. (2006). *Scientific Models-A User-oriented Approach to the Integration of Scientific Data and Digital Libraries*. Retrieved January 14, 2010, from http://www.valaconf.org.au/vala2006/papers2006/55_Hunter_Final.pdf.
- Hussin, S. (n.d). *Tatabahasa Kasus*. Retrieved December 23, 2010, from <http://www.iptho.edu.my/jbm/text/il-02.pdf>.
- Informedness, markedness & correlation. *Journal of Machine Learning Technologies*. 2 (1). 37-63.
- Jaafar, J. (2008). Cabaran dan proses pemantapan bahasa melayu. *Jurnal Pengajian Melayu*, 19, 68-90.

- Jabar, H. Y., & Tengku Mohd, T., S. (2006). Design and implement an automatic Neural Tagger Based language for NLP applications. *Asian Journal of Information Technology*. 5(7). 784-789.
- Jakubicek, M. (2012). *Rule-Based Parsing of Morphologically Rich Languages*. PhD Thesis. Masaryk University.
- Johnson, B., & Shneiderman, B. (1991). Tree-Maps: A Space-Filling Approach to the Visualization of Hierarchical Information. *Proceedings of Visualization 1991*, 284-291. doi: 10.1109/VISUAL.1991.175815
- Jubilado, R. C. (2010). *Clause Structure of Malay in the Minimalist Program*. PhD Thesis. Universiti Kebangsaan Malaysia: Bangi.
- Juhaida, A. B., Khairuddin, O., Mohammad Faidzul, N., & Mohd Zamri, M. (2016). Nuwt: Jawi-Specific Buckwalter Corpus for Malay Word Tokenization. *Journal of ICT*. 15(1). 107-131.
- Jurafsky, D., & Martim, J. H. (2009). *Speech and language processing, An introduction to Natural Language Processing, Computational linguistics, and Speech Recognition*. 2nd Edition. 479-576. Pearson Prentice Hall: New Jersey.
- Kaidi, Z. (2000). *Data visualization*. Retrieved December 15, 2010, from http://www.cs.uic.edu/~kzhao/Papers/00_course_Data_visualization.pdf
- Kakkonen, T. (2007). *Framework and resources for natural language parser evaluation*. PhD Thesis. University of Joensuu, Finland.
- Kementerian Pendidikan Malaysia. (2003). *Kurikulum bersepadu sekolah rendah, Sukatan pelajaran, Bahasa Melayu*. Retrieved January 15, 2010, from http://www.moe.gov.my/bpk/sp_hsp/bm/kbsr/sp_bm_kbsr.pdf
- Kikas, T., & Treumuth, M. (2007). *Automatic Parser Evaluation*. Retrieved December 16, 2014, from <http://math.ut.ee/~treumuth/NLP/syntax.pdf>
- Klein, D., & Manning, C. D. (2003). *Accurate unlexicalized parsing*. Proceedings of the 41st Annual Meeting on Association for Computational Linguistics. July 07-12, Sapporo, Japan. 423-430
- Kong, T. E. (1994). *Natural language analysis in machine translation (MT) based on the string-tree correspondence grammar (STCG)*. Ph.D. thesis., Universiti Sains Malaysia, Penang, Malaysia.
- Kovar, V. (2014). *Automatic Syntactic Analysis for Real-World Applications*. PhD Thesis. Masaryk University, Brno, Czech Republic.
- Laventhal, L., & Barnes, J. (2008). *Usability engineering: process, product, and examples*. Pearson Prentice Hall: United States.
- Lee, B. (2006). *Interactive visualizations for trees and graphs*. PhD Thesis, Universiti Maryland.
- Luboschik, M., & Schumann, H. (2007). Explode to Explain-Illustrative Information Visualization. *11th International Conference Information Visualization (IV'07)*, 301-307. doi: 10.1109/IV.2007.50
- Lund, A.M. (2001). *Measuring usability with the USE Questionnaire*. Retrieved from <http://garyperlman.com/quest/quest.cgi?form=USE>

- Maisarah, Y. (2013). *MYPARSER: A Malay text categorization toolkit using inference rule*. Master Thesis. Universiti Tun Hussein Onn Malaysia, Johor.
- Mazura, M. (2002). *Frasa nama Bahasa Malaysia dan Bahasa Inggeris di dalam teks terjemahan: satu kajian perbandingan struktur, makna dan strategi terjemahan*. Master Thesis. USM.
- Manning, C. D., & Schutze, H. (2000). *Foundations of statistical natural language processing*. 3rd edition. 534. The MIT Press Cambridge, Massachusetts: London, England.
- Mitkov, R. (2004). *The Oxford Handbook of Computational Linguistics*. Oxford University Press: New York.
- Mohd Juzaidin, A. A. (2008). *Pola grammar for automated marking of Malay short answer essay-type examination*. PhD Thesis. Universiti Putra Malaysia: Serdang.
- Mohd Juzaidin, A. A. (2007). Pengkomputeran Linguistik Bahasa Malaysia. *Prosiding Persidangan Kebangsaan Sains Pengaturcaraan 2007: Memacu Penyelidikan Pengaturcaraan ke Arah Masa Hadapan*, 69-76.
- Mohd Juzaidin, A. A., Fatimah, A., Abdul Azim, A. G., & Ramlah, M. (2008). *Identify Malay Sentence Similarity based on Pola Grammar Algorithm*. 12th WSEAS International Conference on Computers. Heraklion, Greece, July 23-25.
- Mooney, R. J. (2004). *Machine learning*. In *The Oxford Handbook of Computational Linguistics*. Editors: Mitkov, R. 376-394. Oxford University Press: New York.
- Muhamad Shahbani, A. B. (2012). *Model reka bentuk konseptual operasian storan data bagi aplikasi kepintaran perniagaan*. PhD Thesis. Universiti Utara Malaysia, Kedah.
- Murugesan, A., & Cassimatis, N. (2006). A model of syntactic parsing based on domain-general cognitive mechanisms. *Prosiding 8th annual conference of the Cognitive Science Society*. Vancouver Canada.
- Musthofa (2010). Computational Linguistics: Model Baru Kajian Linguistik dalam Perspektif Komputer. *Adabiyāt: Jurnal Bahasa dan Sastra*, 9(2).247-271.
- Nasrun, A. (1994). *Verb phrase in Malay: an approach based X-bar theory*. Master Thesis. Universiti Kebangsaan Malaysia: Bangi.
- Nawi, I. (2003). *Budaya Bangau Oh Bangau dalam Bahasa Melayu*. Retrieved January 18, 2010, from <http://www.oocities.com/pendidikmy/berita/berita42003.html>
- Nazri, A. B., Muhammad, S., Shamsinah, S., Norizillah, M. R., & Fatahiyah, M. I. (2006). *Penggunaan komputer dalam pengajaran bahasa*. Retrieved December 28, 2010, from <http://202.28.66.7/smuhammad/pdf/Penggunaan%20Komputer%20dlm%20pengajaran%20bahasa.pdf>
- Nederhof, M. J., & Satta, G. (2013). *Teory of Parsing*. In *The Handbook of Computational Linguistics and Natural Language Processing*. Editors: Clark, E., Fox, C., & Lappin, S. 271-296. Blackwell Publishing Ltd: United Kingdom.
- Nelson, C. M., Punch, R.E., & Donaldson, J.E. (2011). *An Interactive Software Tool for Parsing English Sentences*. Proceedings of the Midstates Conference on Undergraduate Research in Computer Science and Mathematics.

- Nguyen, Q. V., & Huang, M. L. (2002). A Space-Optimized Tree Visualization. *Proceedings of the IEEE Symposium on Information Visualization 2002 (InfoVis'02)*, 85-92. doi: 10.1109/INFVIS.2002.1173152
- Nielsen, J. (1993). *Usability Engineering*. AP Professional: New York.
- Nielsen, J. (2000). *Why you only need to test with five users*. Retrieved from <https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>
- Nik Hassan Basri, N. A. K. (2009). *Teori Bahasa, Implikasinya terhadap pengajaran tatabahasa*. Edisi kedua. Universiti Pendidikan Sultan Idris: Perak.
- Nik Safiah, K. (1995). *Malay Grammar for Academics and Professionals*. Dewan Bahasa dan Pustaka: Kuala Lumpur.
- Nik Safiah, K., Farid M. O., Hashim, H. M., & Abdul Hamid, M. (2009). *Tatabahasa dewan edisi ketiga*. Dewan Bahasa dan Pustaka: Kuala Lumpur.
- Nik Safiah, K., Farid M. O., Hashim, H. M., & Abdul Hamid, M. (2004). *Tatabahasa Dewan Edisi Baharu*. Dewan Bahasa dan Pustaka: Kuala Lumpur.
- Noor Hafhizah, A. R. (2011). *A statistical parser to reduce structural ambiguity in Malay grammar rules*. Thesis Master. Universiti Malaya, Kuala Lumpur.
- Nor Hashimah, J., Junaini, K., & Zaharani, A. (2010). Sosiokognitif pelajar remaja terhadap Bahasa Melayu. *GEMA Online™ Journal of Language Studies*, 10(3), 67-87.
- Norshuhani, Z., & Arina, G. (2010). A Hybrid Approach for Malay Text Summarizer. *Proceedings of the International Multi-Conference on Engineering and Technological Innovation (IMETI 2010)*. June 29-July 02, Florida, USA.
- Norshuhada, S., & Shahizan, H. (2010). *Design Research in Software Development: Constructing and Linking Research Questions, Objectives, Methods and Outcomes*. Sintok: Penerbit Universiti Utara Malaysia.
- Park, Y.U., & Kwon, H.C. (2008). *Korean Syntactic Analysis using Dependency Rules and Segmentation*. International Conference on Advanced Language Processing and Web Information Technology. 59-63.
- Pathiah, A. S. (2012). *A common modeling language for model checkers*. PhD Thesis. Universiti Kebangsaan Malaysia: Bangi.
- Peters, M. (2008). The Development of a Semantic Model for learning Mathematics. *Proceedings of the British Society for Research into Learning Mathematics*. Retrieved January 15, 2015, from <http://www.bsrlm.org.uk/IPs/ip28-2/BSRLM-IP-28-2-14.pdf>
- Phang, S. W., & Zarina, S. (2012). *Asas teknik pemprosesan bahasa*. Universiti Kebangsaan Malaysia: Bangi Selangor.
- Plass, J.L., Moreno, R., & Brunken, R. (2010). *Introduction*. In Cognitive load theory. Cambridge University Press: United States.
- Potemkin, S.B. (2009). *Unsupervised Parsing of the Russian Sentence*. Proceedings of the SENSE Workshop on conceptual Structures for Extracting Natural language SEMantics Moscow, Russia, July.

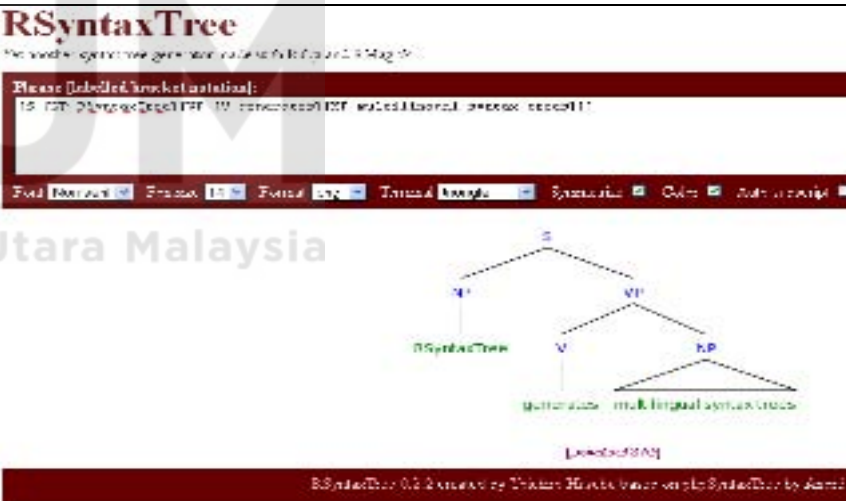
- Powers, D. M. W. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation. *Journal of Machine Learning Technologies*, 2(1), 37-63.
- Prat, N., Comyn-Wattiau, I., & Akoka, J. (2014). Artifact evaluation in information systems design-science research-a holistic view. *The 18th Pacific Asia Conference on Information Systems (PACIS 2014) proceedings*.
- Resnik, P., & Lin, J. (2013). *Evaluation of NLP Systems*. In *The Handbook of Computational Linguistics and Natural Language Processing*. Editors: Clark, E., Fox, C., & Lappin, S. 271-296. Blackwell Publishing Ltd: United Kingdom.
- Ramli, S. (1995). *Sintaksis Bahasa Melayu Penerapan Teori Kuasaan dan Tambatan*. Dewan Bahasa dan Pustaka: Kuala Lumpur.
- Rosmah, A. L. (1995). *Penyemak Sintaksis Ayat Bahasa Malaysia*. (Tesis sarjana). Universiti Kebangsaan Malaysia, Bangi.
- Rozana, K., Nurul Atiqah, A., Eliza Mazmee, M., & Saipunidzam, M. (2011). Malay Language Sentence Checker. *World Applied Sciences Journal 12 (Special Issues on Computer Application & Knowledge Management)*. 19-25.
- Rubin, J., & Chisnell, D. (2008). *Handbook of usability testing: how to plan, design, and conduct effective tests*. 2nd Ed. Wiley Publishing Inc.: United States.
- Rusu, A., Santiago, C., & Jianu, R. (2007). Real-time Interactive Visualization of Information Hierarchies. *11th International Conference Information Visualization (IV'07)*, 117-123. doi: 10.1109/IV.2007.92
- Scheaffer, R. L., Mendenhall III, W., & Ott, R.L. (2006). *Elementary survey sampling*. Sixth Ed. Thomson Brooks/Cole: United States.
- Sekretariat Pusat Majlis Bahasa Melayu IPT Nusantara (2013). *Laporan projek penyelidikan kajian asas kedudukan Bahasa Melayu di Institusi Pengajian Tinggi Awam Malaysia*. Penerbit UMT: Terengganu Malaysia.
- Shalan, K., Farouk, A., Rafea, A. (1999). *Towards an Arabic Parser for Modern Scientific Text*. In *Proceeding of the 2nd Conference on Language Engineering, Egyptian Society of Language Engineering (ELSE)*, Egypt. 103-114.
- Shatnawi, M., & Belkhouche, B. (2012). *Parse Trees of Arabic Sentences Using the Natural Language Toolkit*. The 13th International Arab Conference on Information Technology (ACIT'2012). December 10-13, Kurah Lebanon
- Siti Hajar, A. A. (2009). *Bahasa Melayu II*. Oxford Fajar Sdn. Bhd.: Selangor
- Siti Hajar, A. A. (2011). *Bahasa Melayu I*. Edisi kedua. Oxford Fajar Sdn. Bhd.: Selangor.
- Sleator, D., & Temperley, D. (1991). *Parsing English with a Link Grammar*. Retrieved December 15, 2010, from http://arxiv.org/PS_cache/cmp-lg/pdf/9508/9508004v1.pdf
- Sleator, D., & Temperley, D. (1993). Parsing English with a link grammar. *Proceedings of the Third Annual Workshop on Parsing Technologies*. 1-14. Retrieved December 29, 2010, from <http://www.cs.cmu.edu/afs/cs.cmu.edu/project/link/pub/www/papers/ps/LG-IWPT93.pdf>




- Somuncuoglu, Y., & Yildirim, A. (1999). "Relationship between achievement goal orientation and use of learning strategies". *The Journal of Educational Research*, 92(5), 267-277.
- Soricut, R., & Marcu, D. (2003). *Sentence Level Discourse Parsing using Syntactic and Lexical Information*. Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology. 149-156.
- Spence, R. (2007). *Information visualization, design for interaction*. Pearson Education Limited: England.
- Sun, S. L., Zaidatun, T., & Jamalludin, H. (2007). *Penghasilan modul pembelajaran berasaskan teori beban kognitif untuk subjek teknologi maklumat dan komunikasi*. 1st International Malaysian Educational Technology Convention. Johor Bahru. 1204-1213.
- Suzaimah, R. (2002). *Reka bentuk dan implementasi suatu penghurai bahasa Melayu menggunakan sistem logik selari*. (Tesis sarjana) Universiti Putra Malaysia, Selangor.
- Syamsul Bahrin, Z. (2011). *Mobile game-based learning (MGBL) engineering model*. PhD Thesis. Universiti Utara Malaysia, Kedah.
- Tan, T. S., & Sh-Hussain. (2009). Corpus Design for Malay Corpus-based Speech Synthesis System. *American Journal of Applied Sciences*. 6(4): 696-702.
- Tayal, M.A., Raghuwanshi, M.M., & Malik, L. (2014). *Syntax Parsing: Implementation using Grammar-Rules for English Language*. 2014 International Conference on Electronic Systems, Signal Processing and Computing Technologies. 376-381.
- Thant, W. W., Htwe, T. M., & Thein, N. L. (2012). *Parsing of Myanmar sentences with function tagging*. arXiv preprint arXiv:1205.1603.
- Tullis, T., & Albert, B. (2008). *Measuring the user experience: collecting, analyzing, and presenting usability metrics*. Morgan Kaufmann: USA.
- Tullis, T., & Albert, B. (2013). *Measuring the user experience: collecting, analyzing, and presenting usability metrics (2nd Ed.)*. Morgan Kaufmann: USA.
- University of Maryland. (2003). *Treemap*. Retrieved December 28, 2010, from <http://www.cs.umd.edu/hcil/treemap/>
- Vaishnavi, V. K., & Kuechler, W. (2008). *Design Science Research Methods and Patterns: Innovating Information and Communication Technology*: Auerbach Publications, Taylor & Francis Group.
- Voloshin, V. (2009). *Introduction to graph theory*. Nova Science Publishers, Inc: New York.
- Wang, W., Wang, H., Dai, G., & Wang, H. (2006). Visualization of Large Hierarchical Data by Circle Packing. *Proceedings of the SIGCHI conference on Human Factors in computing systems (CHI '06)*, 517-520. doi: 10.1145/1124772.1124851
- Ware, C. (2000). *Information Visualization: perception for design*. Morgan Kaufmann Publishers: USA.

- Ware, C. (2013). *Information Visualization, perception for design* (Third Edition). Morgan Kaufmann: USA.
- Yoichiro, H. (2012). *RSyntaxTree*. Retrieved April 08, 2012, from <http://yohasebe.com/rsyntaxtree/>
- Yuni Dwi, A. (2005). *Algoritma*. Retrieved January 13, 2011, from http://yuni_dwi.staff.gunadarma.ac.id/Downloads/files/12675/Bab+5+-+Algoritma.pdf
- Zaharani, A., & Nor Hashimah, J. (2012). Incorporating structural diversity in the Malay grammar. *GEMA Online™ Journal of Language Studies*, 12(1), Special Section, 17-34.
- Zaharin, Y. (1986). *Strategies and heuristics in the analysis of a natural language in machine translation*. Ph.D. thesis, Universiti Sains Malaysia, Penang, Malaysia.
- Zaharin, Y. (1998). *Cintailah bahasa kita, suatu tanggapan linguistik berkomputer*. Universiti Sains Malaysia: Pulau Pinang, Malaysia.
- Zaharin, Y. (2000). Computational linguistics in Malaysia. *Proceedings of the 38th Annual Meeting on Association for Computational Linguistics*. Hong Kong. Pages: 1-2.
- Zaini, A., Mohmad Noor, M. T., Ikhsan, O., Norila, M. S., Abu Bakar, Y., & Abdul Talib, M. H. (2012). *Perkembangan pendidikan di Malaysia: Falsafah dan dasar* (KPF 3012), panduan kursus Program Ijazah Sarjana Muda Pendidikan UPSI, 56.
- Zhao, J., Chevalier, F., Collins, C., & Balakrishnan, R. (2012). Facilitating Discourse Analysis with Interactive Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 18(12). 2639-2648.
- Zulkifley, H. (2012). *Bahasa Melayu bahasa universal*. Pendidikan bahasa melayu di Malaysia: Pemilihan dan implikasi (Bab 10). Universiti Kebangsaan Malaysia: Selangor.
- Zuraidah, M. D. (2010). Processing natural Malay texts: A data-driven approach. *TRAMES: A Journal of the Humanities & Social Sciences*, 14(1), 90.

Lampiran A

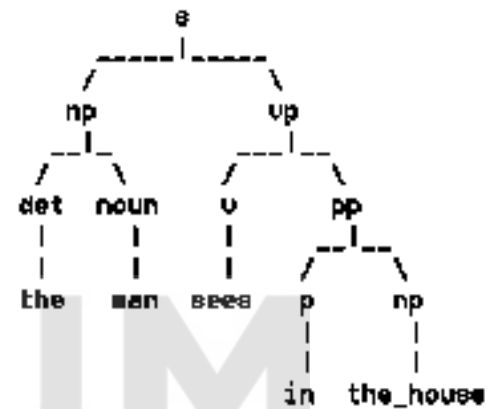
Aplikasi pohon sintaksis untuk BI

Aplikasi	Tahun	Kaedah input	Kelemahan	Implikasi untuk pohon sintaksis BM	Contoh pohon sintaksis / antara muka
RSyntaxTree ree	2009/2 010	Simbol kurungan	Sukar bagi pengguna yang tidak memahami format penulisan input berbentuk simbol kurungan (Berdasarkan kepada phpSyntaxTree)	Paparan antara muka boleh dirujuk untuk reka bentuk antara muka pakej gabungan VPS.	 <p>The screenshot shows the RSyntaxTree application window. At the top, it displays the title 'RSyntaxTree' and a subtitle 'No need to generate grammar rules manually found in Mag 09'. Below this, there is a text input field containing the sentence 'RSyntaxTree menggunakan multiple syntax trees'. A syntax tree diagram is generated below the input, showing a root node 'S' branching into 'NP' and 'VP'. The 'NP' node branches into 'RSyntaxTree', and the 'VP' node branches into 'V' (using) and 'NP' (multiple syntax trees). The diagram is rendered in green text on a white background. At the bottom of the window, there is a footer that reads 'RSyntaxTree 0.2.2 created by Chuan Hwa Lee based on phpSyntaxTree by Jami'.</p>

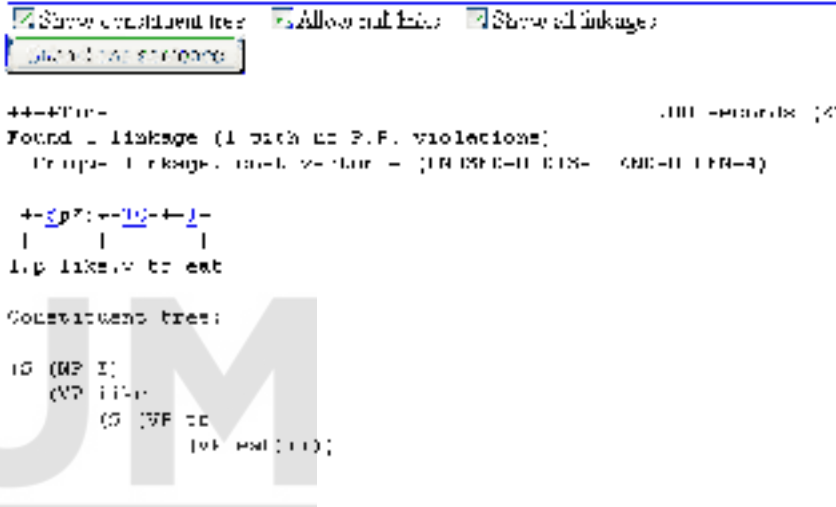
SynView	2009	Simbol	Memerlukan perisian	Susunan nod dan anak	
		kurungan	LaTex dan penganalisis	panah yang kemas	
		menggung	luaran	boleh dirujuk walaupun	
		akan		aplikasi ini	
		notepad		menggunakan kaedah	
				penghurai bawah-atas.	
phpSinta	2003	Simbol	Sukar bagi pengguna yang	Paparannya antara muka	
kTree		kurungan	tidak memahami penulisan	boleh dirujuk untuk	
			berbentuk simbol	reka bentuk antara	
			kurungan	muka pakej gabungan	
				VPS.	

Lehner's prolog tree drawing	1994	Simbol Prolog	Menyukarkan pengguna yang tidak memahami struktur prolog	Kaedah lakaran pembahagian subjek- predikat boleh dirujuk.
---------------------------------------	------	------------------	--	--

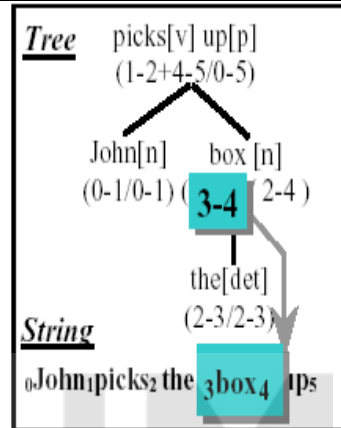
SWI Prolog Question 1.4



UUM
Universiti Utara Malaysia

Link Grammar (LG)	1991	Berbentu k ayat	<p>-Ayat tidak dikategorikan kepada subjek dan predikat</p> <p>-LG adalah aplikasi yang menjalankan kajian tentang perkaitan antara tatabahasa yang terdapat dalam ayat dan bukanlah aplikasi pohon sintaksis yang berbentuk hierarki atau ciri-ciri lain yang diperlukan dalam kajian ini, tetapi masih boleh dijadikan rujukan ekoran LG juga sebuah aplikasi untuk menganalisis ayat dan kaedah yang digunakan adalah kaedah asas dalam menganalisis ayat bahasa tabii.</p>	<p>Teknik menganalisis ayat yang dilakukan iaitu</p> <ol style="list-style-type: none"> 1) Membaca setiap perkataan 2) Buat padanan dengan pangkalan data bagi setiap perkataan 3) Padanan dengan rumus 4) Visualisasi <p>Kaedah analisis ayat yang dilakukan sesuai dijadikan rujukan kerana kaedah yang digunakan diaplikasikan untuk menganalisis ayat dan penggunaan rumus</p>	<p>You can go back to the Link Grammar's front page</p> 
-------------------------	------	--------------------	--	--	---

SSTC (Structure-string tree correspondence)	1998	Ayat	<p>Banyak turutan yang terlibat seperti pengecaman perkataan berasaskan contoh yang diberi dalam pangkalan data, pembahagian kepada sub-pohon sintaksis, dan menggunakan simbol yang dinamakan SSTC bagi setiap perkataan</p> <p>-Dibangunkan untuk mesin terjemahan yang menghasilkan pohon sintak untuk kedua-dua bahasa serentak.</p> <p>-Tiada pembahagian kepada subjek dan predikat serta tiada rumus yang digunakan.</p>	<p>Model yang digunakan boleh dirujuk untuk mendapatkan gambaran kaedah pemprosesan ayat yang digunakan.</p>
--	------	------	---	--



Lampiran B

Contoh surat persetujuan responden

SURAT PERSETUJUAN RESPONDEN

Tujuan: Soal selidik untuk pengguna prototaip pohon sintaksis pakej gabungan VPS

Saya Muhammad Imran bin Abdul Karim No. kad pengenalan 030413-02-0144-7 dengan ini bersetuju untuk mengambil bahagian secara sukarela dalam soal selidik yang dijalankan oleh pengkaji Yusnita binti Muhamad Noor bernombor metrik 92715 daripada Pusat Pengajian Pengkomputeran, Universiti Utara Malaysia untuk pengguna prototaip pohon sintaksis pakej gabungan VPS.

Saya telah diberi penjelasan secara menyeluruh mengenai tujuan kajian dan soal selidik yang dilakukan termasuklah diberi kebenaran untuk menggunakan prototaip yang disediakan. Saya juga mendapat penerangan mengenai tujuan ini secara bertulis di belasan soal selidik. Saya memahami bahawa saya berhak menarik diri jika tidak bersetuju untuk menyertai sebagai responden pada bila-bila masa tanpa memberi sebarang alasan. Saya juga memahami bahawa sebarang maklumat yang berkaitan identiti saya akan dirahsiakan.

Saya bersetuju/tidak bersetuju untuk imej/gambar/rakaman video/rakaman suara digunakan dalam apa jua bentuk penerbitan atau pembentangan.

*potong yang tidak berkenaan

Tandatangan: <u>[Signature]</u> (responden)	Tandatangan: <u>[Signature]</u> (saksi)
Tarikh: <u>01/08/2016</u>	Nama: <u>Muhammad Imran bin Abdul Karim</u> No. K/P: <u>03102402-0144</u>

Saya mengesahkan bahawa saya telah menerangkan kepada responden ini tujuan kajian dan tujuan soal selidik dilakukan.

Tandatangan: [Signature]
(pengkaji)

Tarikh: 01/08/2016

Lampiran C

Surat persetujuan pengetua

Pengetua
Sek. Men. Keb. Chunglun
06010 Jitra
Kedah

SESILAPA YANG BERKENAAN 1 Ogos 2016

Tuan / Puan

KEBENARAN MENJALANKAN SOAL SELIDIK

Saya Omara b. Baharom sebagai pengetua Sekolah Menengah Kebangsaan Chunglun, dengan ini membenarkan pelajar bernama Yusra hini Muhammad Noor daripada UUM untuk menjalankan soal selidik terhadap 30 orang pelajar tingkatan satu hingga tingkatan lima pada tarikh 1 Ogos 2016 bertempat di antara pelajar.

Sekian, terima kasih.



(**OMARA BIN BAHAROM S.C.K**)
Pengetua
Sekolah Menengah Kebangsaan Chunglun
06010 Jitra, Kedah Darul Aman.

Universiti Utara Malaysia

Lampiran D

Surat kebenaran pengumpulan data

 **AWANG HAD SALLEH**
GRADUATE SCHOOL OF ARTS AND SCIENCES
UUM College of Arts and Sciences
Universiti Utara Malaysia
06010 UUM SINTOK
KEDAH DARUL AMAN
MALAYSIA

 **UUM**
Universiti Utara Malaysia

Tel: 04-926 5499/0495271
Faks (Fax): 04-926 5227/5226
Laman Web (Web): <http://www.uum.edu.my>

KEDAH AMAN MAKMUR - BERSAMA MENJALU TRANSFORMASI

UUM/CAS/AHSGS/92715

31 July 2016

TO WHOM IT MAY CONCERN

Dear Sir/Madam

DATA COLLECTION FOR PROJECT PAPER/ THESIS

This is to certify that Yuxnita binti Muhamad Noor (matric number: 92715) is a full time postgraduate student in Doctor of Philosophy (Information Technology) at UUM College of Arts and Sciences.

She needs to do her field study and data collection for her project paper/thesis in order to fulfill the partial requirements of her graduate studies.

We sincerely hope that your organization will be able to assist her in the data collection and the distribution of the questionnaires for her research. All the information collected only for academic purposes and will be kept confidential.

Thank you.

"KNOWLEDGE, VIRTUE, SERVICE"

Yours faithfully



WAS NORFASHMA BINTI WAN MIN
Senior Assistant Registrar
for Dean
Awang Had Salleh Graduate School of Arts and Sciences
UUM College of Arts and Sciences

Universiti di Rantau Hijau - The University in a Green Forest



Lampiran E

Instrumen penilaian pakar

Tuan/Puan

PENILAIAN PAKAR MODEL PAKEJ GABUNGAN VPS

Saya Yusnita binti Muhamad Noor, pelajar PhD dalam bidang Teknologi Maklumat di Universiti Utara Malaysia. Kajian PhD saya bermatlamat utama untuk menghasilkan model dan algoritma pakej gabungan VPS iaitu gabungan antara semakan ayat, cadangan pembetulan ayat, visualisasi pohon sintaksis (VPS) dan set atribut perkataan (kelas kata, kata terbitan, terjemahan, imej dan contoh ayat).

Penentusahan model adalah salah satu sub-objektif yang perlu dicapai dalam kajian ini. Soalan yang diajukan adalah berdasarkan kriteria pengesahan komponen model seperti yang terdapat dalam borang penilaian. Diharapkan agar Tuan/Puan sudi meluangkan masa untuk menjawab soalan yang diberikan. Kerjasama yang diberikan sangat dihargai. Sebarang pertanyaan boleh hubungi saya di alamat e-mel (eita_yusnita84@yahoo.com).

Terima kasih atas bantuan dan masa yang diluangkan.

MAKLUMAT PERIBADI

Nama:

Jawatan:

Umur:

Tahap pendidikan tertinggi: _____ bidang: _____

Pengalaman sebagai Munsyi Dewan/pakar bidang: ___ tahun

SOALAN PENILAIAN TERHADAP MODEL PAKEJ GABUNGAN VPS

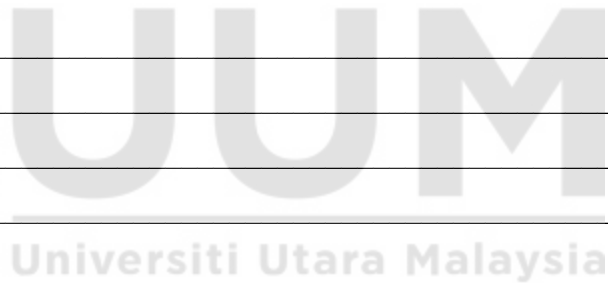
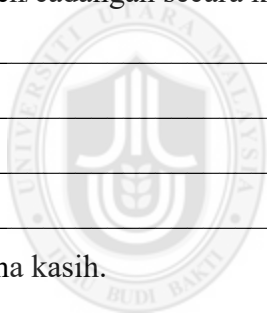
Bil	Kriteria penilaian	Sangat tidak setuju (1)	Tidak setuju (2)	Tidak pasti (3)	Setuju (4)	Sangat setuju (5)
1.	Model ini senang difahami					
2.	Turutan dan proses yang terlibat adalah jelas					
3.	Model ini bersesuaian dengan visualisasi pohon sintaksis (VPS) ayat Bahasa Melayu					
4.	Model ini memberi bantuan pemahaman dan pembelajaran ayat BM dengan paparan struktur frasa, kelas kata dan perkataan dalam bentuk VPS					
5.	Model ini memberi sokongan mendatang terhadap keperluan aplikasi pemprosesan bahasa tabii (NLP)					
6.	Komponen-komponen model adalah bersesuaian untuk membuat VPS ayat BM					
7.	Komponen semakan ayat diperlukan dalam pembelajaran ayat secara VPS					
8.	Komponen cadangan ayat diperlukan dalam VPS untuk memberi cadangan pembetulan ayat mengikut RSF					
9.	Komponen VPS boleh membantu pemahaman struktur ayat, kelas kata dan perkataan					

10.	Set atribut perkataan (kelas kata, kata terbitan, terjemahan, imej, ayat contoh) diperlukan dalam VPS untuk memberi pemahaman yang lebih mendalam tentang struktur ayat dan perkataan serta boleh difahami dalam konteks ayat yang lain.					
-----	--	--	--	--	--	--

Adakah terdapat komponen lain yang dirasakan perlu dimasukkan dalam model ini untuk membuat VPS ayat BM?

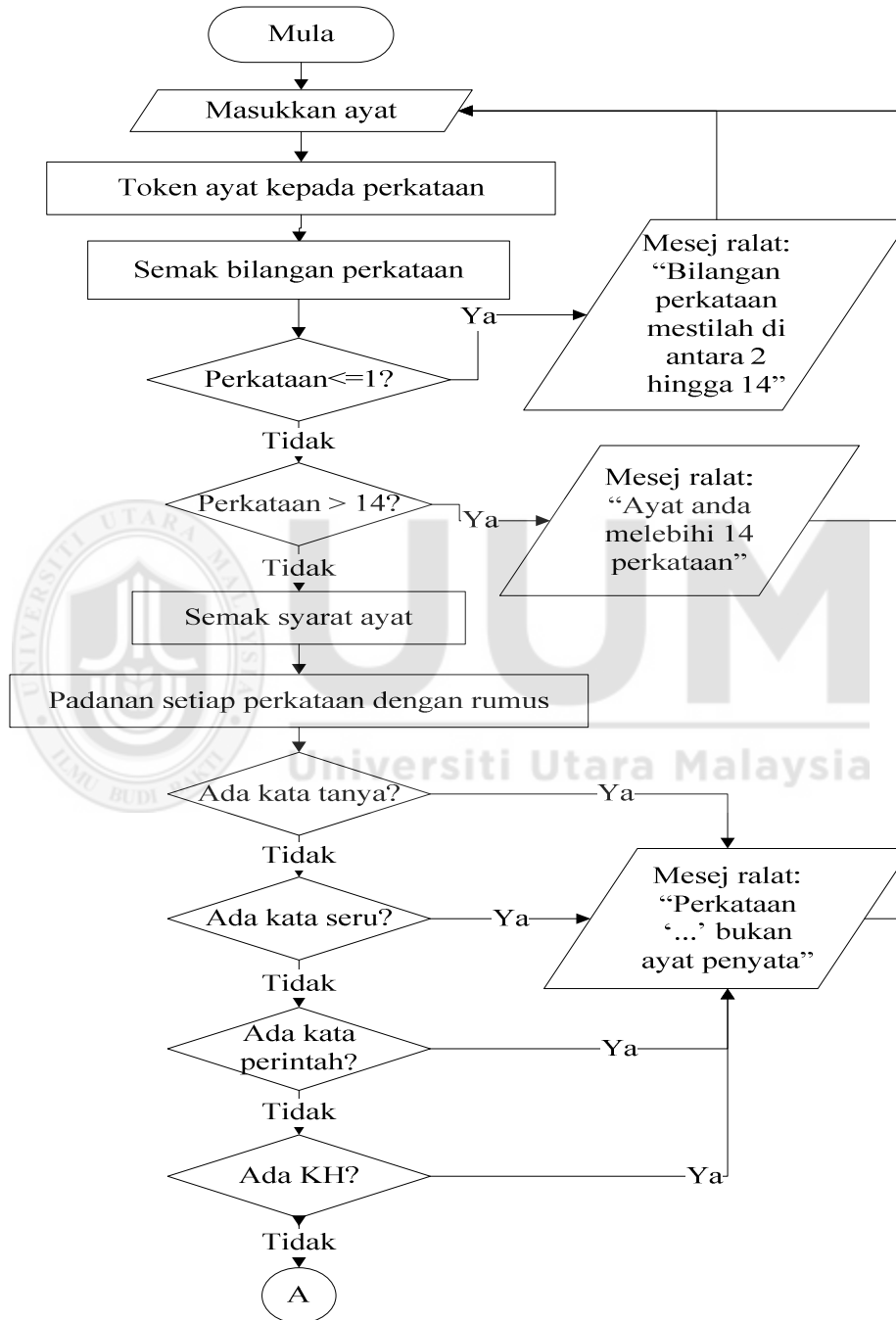
Komen/cadangan secara keseluruhan

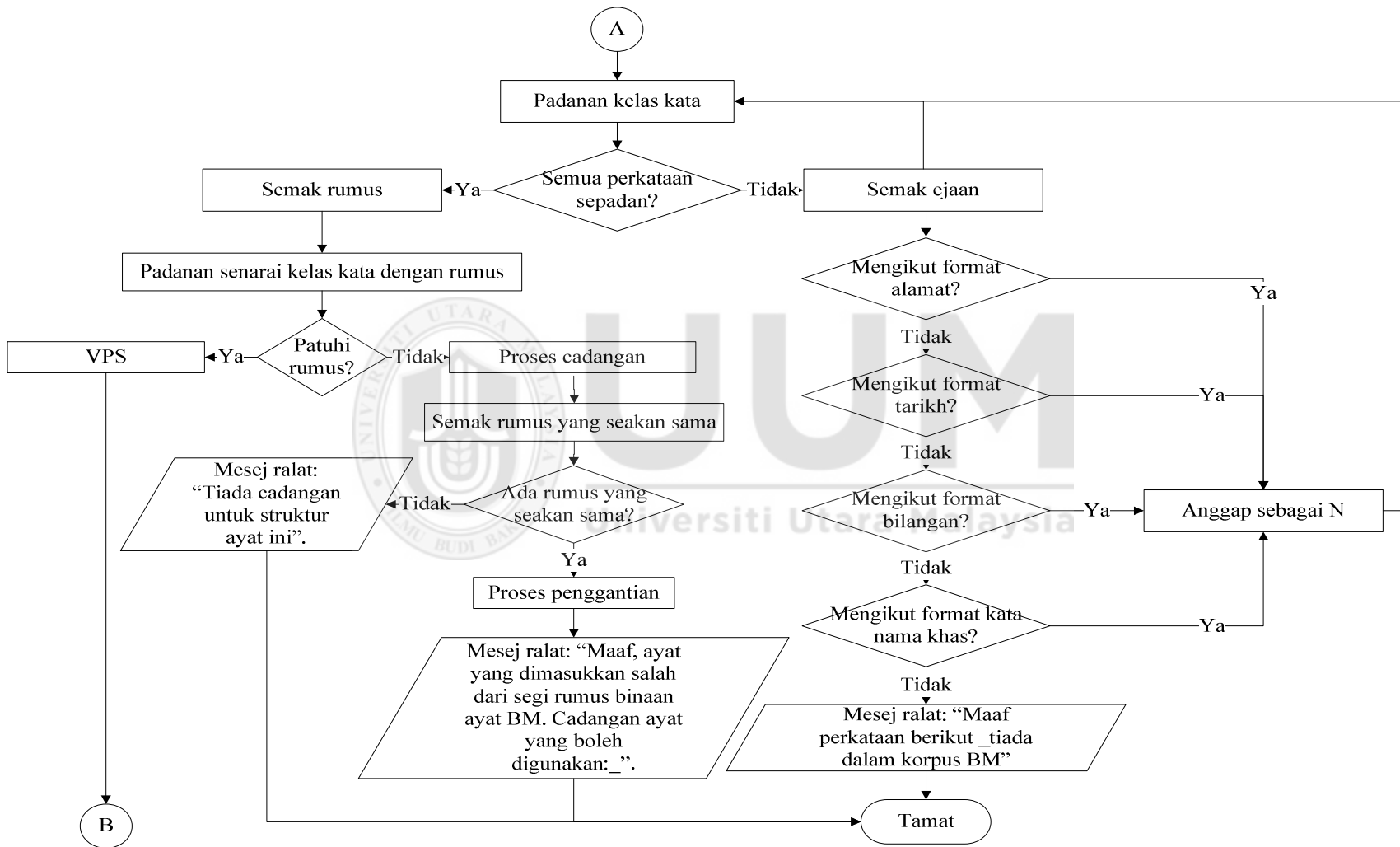
Terima kasih.

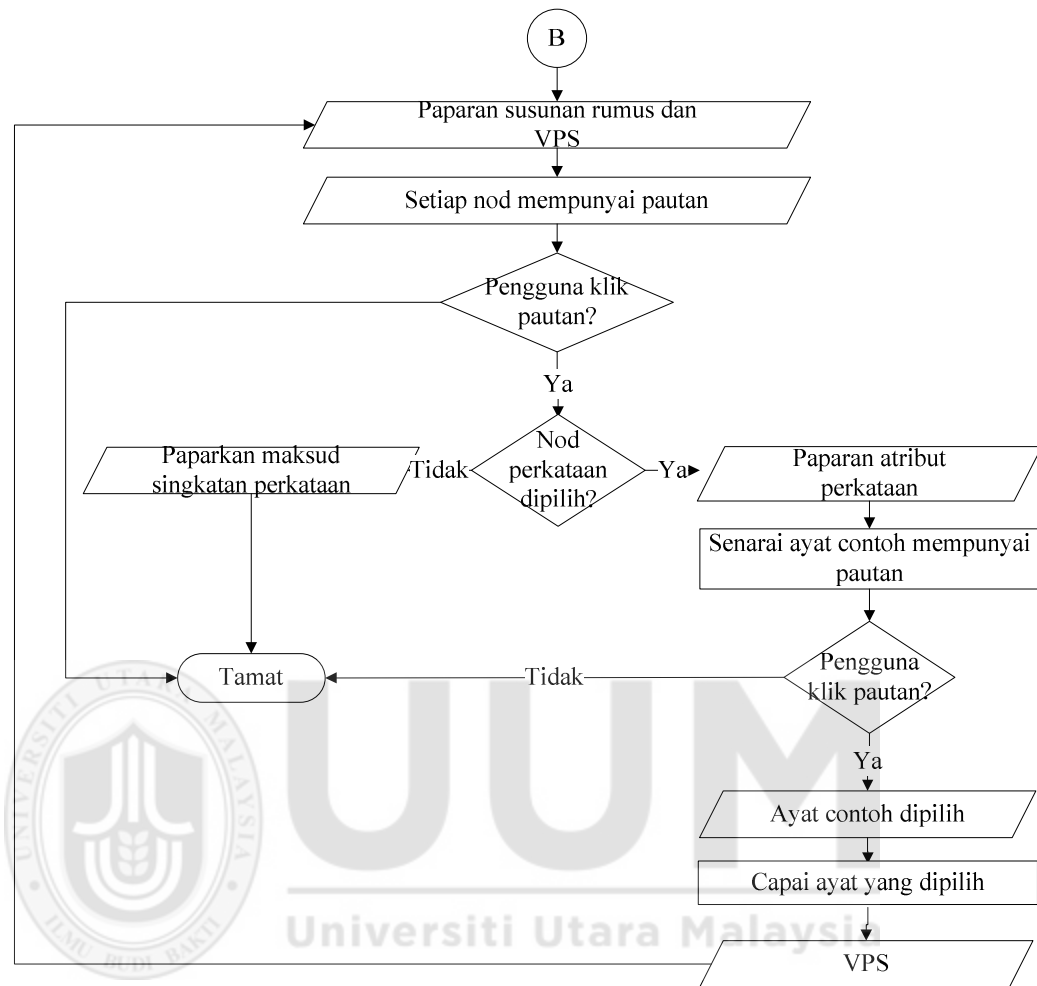


Lampiran F

Carta alir VPS dengan output tambahan







Lampiran G

Rumus *X-bar*

Subjek Frasa Nama (SN^{''})

N	N N	N PK	KBIL N	N N'
N PENT	N KS''	N PA	KBIL N'	N' PA
KAD N	N' PK	N' PENT	N' KAD	N KBIL
KNF N	KPM N	KPM N'	N' KS''	KB N
KAD N'	KNF N'			

Frasa Nama (N^{''})

N	KB N	KB N'	N KB	N' KB
N PENT	PENT N'	N A''	N N	N K''
N KBIL	N KS''	N KAD	KAD N	N' PENT
N'A''	N' K''	N' KS''	N' KBIL	N' N'
PENT	KNF N	KNF N'	N N'	KAD N'
KPM N'	KPM N	KBIL	N' N	KARAH N'
N' KAD	KBIL KS''			

Frasa Nama Pertengahan (N')

N N	N KB	KAD N	KB N	KBIL N
N KBIL	PENT N	N PENT	KNF N	N KNF
N N'	N' KS''	N K	N A	N' K''
N' N'	N A''	KPM N		

Frasa Kerja (K'')

K	KB K	K KB	K' KB	KB K'
KAD K	K KAD	K N''	KP K'	KNF K
K K	K A''	KNF K'	KPN K	KPM K'
KAD K'	K' K	K' N''	KPM K	K' KP
K' K'	K KS''	K' KS''		

Frasa Kerja Pertengahan (K')

KK	KN	KPK	KBK	KADK
KA	KKS''	K'N''	K'KS''	KBK'
KNFK	KK'	K'K	KKAD	

Frasa Adjektif (A'')

A	AKS''	KB A	KB A'	KAD A
A'KAD	AKP	AN''	AK''	KP A
AA	KNF A'	AA'	A'A	KPM A
KPM A'	A'N''	KAD A'	A'K	A'KS''
KP A'	A'KS''	KNF A'	A'K''	KPM A'

Frasa Adjektif Pertengahan (A')

AA	KAD A	A KB	A KAD	KB A
AN''	KP A	AKP	KNF A	AK''

Frasa Sendi Nama (KS'')

KS	KS KS	KS N''	KPM KS'	KB KS'
KAD KS'	KNF KS'	KS' N''		

Frasa Sendi Nama Pertengahan (KS')

KS KS	KAD KS	KS N''	KS KAD
-------	--------	--------	--------

Lampiran H

Ayat uji kaji

Pola ayat frasa nama (N")

1. Pegawai polis itu baik.
2. Keindahan putrajaya ini saya kagumi.
3. Sajak ialah puisi moden.
4. Mereka ialah penyelidik bebas.
5. Perwatakan ialah sifat watak tersebut.
6. Teguran ialah hasil perbuatan menegur.
7. Tema ialah persoalan pokok sesebuah cerita.
8. Plot ialah jalan cerita sesebuah cereka.
9. PDRM ialah Polis Diraja Malaysia.
10. Senarai kumpul namakan ialah satu teknik memperkaya perbendaharaan kata.
11. Beberapa orang budak perempuan itu murid di sekolah saya.
12. Alat muzik adalah sebuah permainan.
13. Engkerurai ialah sejenis alat muzik warisan.
14. Berita tentang kadar kemalangan jalan raya kerap dilaporkan .
15. Kita harus pikul tanggungjawab.
16. Kesannya saya berasa amat sukar bangun pagi.
17. Kata penyambung ayat ialah kata hubung.
18. Temenggung Abu Bakar akhirnya diiktiraf sebagai Sultan johor.

Pola ayat frasa kerja (K")

1. Dia telah berbual-bual dengan rakannya.
2. Syahida selalu menggunakan internet.
3. Masyarakat Sarawak menarikan tarian ini .
4. Dia telah mengajak seorang rakannya .
5. Mereka diajar persediaan untuk memamah.
6. Dia merasakan kata-kata ibunya itu dahulu memang benar.
7. Anda mengikut keluarga bercuti di sebuah tempat peranginan.
8. Penggandaan berentak menggandakan perkataan.
9. Hutan ini menjadi sumber utama bekalan kayu negara.
10. Saya ada membawa beberapa helai kain.
11. Kami berlatih nasyid kontemporari.
12. Fikirannya tidak terganggu lagi.
13. Anda telah mengetahui keperluan makan pelbagai jenis makanan.
14. Negara asing tidak boleh campur tangan.
15. Aku berusaha mencari sesuatu untuk memukul ular itu.
16. Beliau mewujudkan ruangan khas program jawi.
17. Kita masih ada jalan untuk memartabatkan tulisan jawi.
18. Hilmi membalut luka pada tangannya.
19. Anda telah menyediakan borang soal selidik untuk tujuan tersebut.

20. Majlis tersebut diadakan pada waktu pagi hari persekolahan.
21. PKN tidak mahu mengganggu lepasan sekolah melanjutkan pelajaran.
22. Beribu-ribu ekor burung ruak-ruak menemukan sangkar besar itu.
23. Serangan wabak SARS menyebabkan seluruh dunia gempar.
24. Kawasan tersebut juga menjadi habitat pelbagai jenis burung.
25. Imran mencuba kemahiran mencantumkan pokok nanas.
26. Nirmala telah menolak pujukan jurujual.
27. Agama juga menganggap semua orang Islam bersaudara.
28. Saya amat bersimpati terhadap nasib ketiga-tiga beranak tersebut.
29. Bayaran masuk pertandingan boleh dibuat secara tunai.
30. Azli telah menerima sijil perkhidmatan cemerlang daripada majikannya.
31. Pak Ismawan sering terasa Iswan ada di sisinya.
32. Usaha ini menanamkan semangat muhibah.
33. Beliau mengajukan permohonan.
34. Mereka melaksanakan pembedahan itu.
35. Biotek juga menghasilkan haiwan spesies baharu.
36. Saya boleh membawa mereka berurusan dengan kerajaan negeri.
37. Penjaga garisan sudah mengangkat bendera kuning.
38. Kita boleh memperingati hari kemerdekaan dengan pelbagai cara.
39. Khairi telah berkenalan dengan rakan-rakan baharu .
40. Bahan buangan ini dapat dikitar semula.
41. Saya berpendapat kerjaya kepolisan sangat sesuai dengan jiwa saya.
42. Ujian ini dijangka akan dikendalikan oleh kumpulan penilai bertauliah.
43. Datuk Bahaman telah mengambil hasil hutan.
44. Ibu keluar dari kereta.
45. Hari Gawai menjadi perayaan utama di negeri saya.
46. Sarawak juga merupakan destinasi pelancongan utama di negara kita.
47. Syahida menghabiskan masa cuti hujung minggu.
48. Mercu gunung api merupakan satu daripada ciri terkemuka taman ini.
49. Langkah ini memupuk nilai murni dalam kalangan belia.
50. Pelantar menggerudi petroleum didirikan di pantai Laut China Selatan.
51. Usaha ini dapat menarik minat mereka.
52. Tepukan diberi setiap kali peserta membuat persembahan.
53. Kita boleh mendapat maklumat dengan cara menyoal sesuatu pihak.
54. Midun menunjuk ke arah depan.
55. Anda diajak oleh rakan menyaksikan perlumbaan Formula 1.
56. Mereka menggunakan pendatang tanpa izin untuk mengaut keuntungan.
57. Pindahan organ merupakan pencapaian sains perubatan moden.
58. Dia melabuhkan punggungnya di atas lantai .
59. Saya harap anda tidak bermasalah untuk bergaul dengan rakan-rakan.
60. Anda tidak menghadapi sebarang penyakit.
61. Makanan ini menjadi kegemaran pada setiap kali hari raya.
62. Beberapa langkah telah dilaksanakan kerajaan untuk memperkukuh integrasi nasional.
63. Mereka tidak mengambil peduli akan nilai sambutan hari kebangsaan.

64. Rakan anda tidak mahu menggunakan urus niaga perkhidmatan e-dagang.
65. Kepantasan ERL seperti menampakkan baris-baris kelapa sawit berlari .
66. Majlis ceramah kerjaya akan diadakan tidak lama lagi.
67. Pengerusi meminta semua ahli hadir menjayakan projek tersebut.
68. Setiausaha telah mengedarkan kertas kerja.
69. Beliau mengadakan program hari kesihatan.
70. Kami akan membeli kontrak Encik Syukur.
71. Dua orang guru sepenuh masa telah ditugaskan mengajar ahli kelab.
72. Pertubuhan ini dapat membentuk belia mengikut acuan kita sendiri.
73. Sastera tradisional boleh dipermudah untuk bacaan murid sekolah rendah.
74. Kita dapat melihat bintang bertaburan di langit .
75. Para pengunjung dapat melihat keunikan seni reka menara ini.
76. Kekurangan vitamin boleh menyebabkan seseorang itu mengalami rabun malam.
77. Kita akan dapat menghasilkan pelbagai bentuk karya.
78. Seorang lelaki separuh umur masuk.
79. Ruang ini boleh diturunkan untuk memisahkan tempat duduk penonton.
80. Anda diberi peluang menyertai satu pertandingan perbahasan.
81. Kejayaan polis itu melegakan orang ramai.
82. Hujah peguam itu berasaskan fakta kukuh.
83. Keahlian kelab ini terhad.
84. Pelajar lelaki itu tersenyum.
85. Cucu-cucu tidak mahu beredar.
86. Semuanya tidak akan tersangkut.
87. Aku tidak bersekolah hari ini.
88. Singapura tidak menganggotai majlis ini.
89. Bapa tidak pergi ke pekan pagi ini.
90. Anda tidak bersependapat dengan rakan anda.
91. Beliau tidak mengalah.
92. Bangunan Sultan Ahmad Samad merupakan bangunan tinggalan masa lampau.
93. Pelajar Islam mengambil kesempatan bersolat zuhur di masjid itu.
94. Ayah Khairi berbincang dengan keluarganya.
95. Surat tidak rasmi disebut juga sebagai surat kiriman biasa.
96. Syarikat tempat Hadi bekerja menyalahkan Hadi atas kerugian tersebut.
97. Ahli kelab akan diajar tentang sukan ini.
98. Rakan anda memberikan hujahnya tentang simpanan gas asli di negara kita.

Pola ayat frasa adjektif (A")

1. Reka bentuk masjid ini amat menarik.
2. Pemain bola itu amat lincah.
3. Mereka tidak tahu berbahasa Mandarin.
4. Saya tidak suka membuang masa.
5. Saya tidak tahu.
6. Midun tidak takut.

-
7. Anda tidak prihatin.
 8. Hamid sungguh rendah hati.
 9. Ibu sungguh baik.
 10. Klia sungguh istimewa.
 11. Susila amat gembira.
 12. Hilmi sangat gembira.
 13. Hujan amat lebat.
 14. Dia sungguh bijak.
 15. Markahnya paling tinggi.
 16. Lantainya amat bersih.
 17. Kawasanya amat bersih.
 18. Khairi terus melangkah ke arah rakan-rakannya.
 19. Anda ingin mencari maklumat untuk dimuatkan dalam majalah sekolah.
 20. Pelajar perlu lulus dalam setiap peperiksaan.
 21. Usaha perlu dilakukan bagi menggalakkan golongan belia bergiat aktif
 22. Lokasi ini memang strategik sebagai tempat menjamu selera
 23. Projek itu amat penting disiapkan dalam masa dua minggu.
 24. Perpaduan kaum amat penting dalam usaha mengekalkan keamanan.
-

Pola ayat frasa sendi nama (KS")

- | | |
|---|---|
| 1. Anda sebagai pencadang utama. | 7. Anda ke klinik gigi. |
| 2. Udara di situ bersih. | 8. Internet sebagai sumber maklumat. |
| 3. Aliah daripada keluarga sederhana. | 9. Anda sebagai pengarah |
| 4. Gambar di sebelah menunjukkan struktur gigi manusia. | planetarium. |
| 5. Kubah pada bangunan merupakan ciri seni bina Islam. | 10. Mereka di Johor Bahru sekarang. |
| 6. Saya daripada syarikat Agro Ria. | 11. Malaysia sebagai destinasi pelancongan. |

Lampiran I

Hasil uji kaji cadangan pembedulan ayat

Ayat uji kaji	Ayat salah	Cadangan pembedulan ayat
Keindahan Putrajaya ini saya kagumi	kagumi keindahan Putrajaya ini saya	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>Kagumi keindahan putrajaya ini saya</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: saya kagumi keindahan putrajaya ini</p>
Sajak ialah puisi moden	ialah puisi moden sajak	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>ialah puisi moden sajak</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: sajak ialah puisi moden</p>
Mereka ialah penyelidik bebas	Bebas mereka ialah penyelidik	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>bebas mereka ialah penyelidik</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: mereka ialah penyelidik bebas</p>
Perwatakan ialah sifat watak tersebut	ialah sifat watak tersebut perwatakan	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>ialah sifat watak tersebut perwatakan</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: perwatakan ialah sifat watak tersebut</p>

<p>Teguran ialah hasil perbuatan menegur</p>	<p>ialah hasil perbuatan menegur teguran</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan ialah hasil perbuatan menegur teguran Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: teguran ialah hasil perbuatan menegur</p>
<p>Tema ialah persoalan pokok sesebuah cerita</p>	<p>ialah persoalan pokok sesebuah cerita tema</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan ialah persoalan pokok sesebuah cerita tema Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: tema ialah persoalan pokok sesebuah cerita</p>
<p>Plot ialah jalan cerita sesebuah cereka</p>	<p>ialah jalan cerita sesebuah cereka plot</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan ialah jalan cerita sesebuah cereka plot Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: plot ialah jalan cerita sesebuah cereka</p>
<p>PDRM ialah Polis Diraja Malaysia</p>	<p>ialah Polis Diraja Malaysia PDRM</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan ialah Polis Diraja Malaysia pdrm Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: pdrm ialah polis diraja malaysia</p>
<p>Senarai kumpul namakan ialah satu teknik memperkaya perbendaharaan kata</p>	<p>satu teknik memperkaya perbendaharaan kata senarai kumpul namakan ialah</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan ialah satu teknik memperkaya perbendaharaan kata senarai kumpul namakan ialah Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: teknik memperkaya perbendaharaan kata ialah senarai kumpul namakan</p>

Alat muzik adalah sebuah permainan	adalah sebuah permainan alat muzik	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>sedia untuk serangan kali ini saya mahu kamu semua bersiap</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: saya mahu kamu semua bersiap sedia untuk serangan kali ini</p>
Kata penyambung ayat ialah kata hubung	ialah kata hubung kata penyambung ayat	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>ialah kata hubung kata penyambung ayat</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: ayat ialah kata hubung kata penyambung</p>
Dia telah mengajak seorang rakannya	mengajak seorang rakannya dia telah	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>mengajak seorang rakannya dia telah</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: dia telah mengajak seorang rakannya</p>
Mereka diajar tentang persediaan untuk memanah	diajar tentang persediaan untuk memanah mereka	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>diajar tentang persediaan untuk memanah mereka</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: diajar tentang persediaan untuk memanah mereka</p>
Dia merasakan kata-kata ibunya itu dahulu memang benar.	merasakan kata-kata ibunya itu dahulu memang benar dia	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>dia merasakan kata-kata ibunya itu dahulu memang benar dia</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: dia merasakan kata-kata ibunya itu dahulu memang benar dia</p>

<p>Anda mengikut keluarga bercuti di sebuah tempat peranginan.</p>	<p>mengikut keluarga bercuti di sebuah tempat peranginan anda</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan ini menjadi sumber utama bekalan kayu-kayan negara hutan Pohon Penghurai Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM. Cadangan ayat yang boleh digunakan: hutan ini menjadi sumber utama bekalan kayu-kayan negara</p>
<p>Hutan ini menjadi sumber utama bekalan kayu-kayan negara.</p>	<p>ini menjadi sumber utama bekalan kayu-kayan negara hutan</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan ini menjadi sumber utama bekalan kayu-kayan negara hutan Pohon Penghurai Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM. Cadangan ayat yang boleh digunakan: hutan ini menjadi sumber utama bekalan kayu-kayan negara</p>
<p>Anda telah mengetahui keperluan makan pelbagai jenis makanan.</p>	<p>telah mengetahui keperluan makan pelbagai jenis makanan anda</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan telah mengetahui keperluan makan pelbagai jenis makanan anda Pohon Penghurai Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM. Cadangan ayat yang boleh digunakan: telah mengetahui keperluan makan pelbagai jenis makanan</p>
<p>Aku berusaha mencari sesuatu untuk memukul ular itu.</p>	<p>mencari sesuatu untuk memukul ular itu aku berusaha</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan mencari sesuatu untuk memukul ular itu aku berusaha Pohon Penghurai Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM. Cadangan ayat yang boleh digunakan: mencari sesuatu untuk memukul ular itu aku berusaha</p>
<p>Kita masih ada jalan untuk memartabatkan tulisan jawi.</p>	<p>ada jalan untuk memartabatkan tulisan jawi kita masih</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan ada jalan untuk memartabatkan tulisan jawi kita masih Pohon Penghurai Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM. Cadangan ayat yang boleh digunakan: ada jalan untuk memartabatkan tulisan jawi</p>

<p>Hilmi membalut luka pada tangannya</p>	<p>membalut luka pada tangannya Hilmi</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>membalut luka pada tangannya hilmi</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: hilmi membalut luka pada tangannya</p>
<p>Anda telah menyediakan borang soal selidik untuk tujuan tersebut.</p>	<p>telah menyediakan borang soal selidik untuk tujuan tersebut anda</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>Anda telah menyediakan borang soal selidik untuk tujuan tersebut anda</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: anda telah menyediakan borang soal selidik untuk tujuan tersebut</p>
<p>Majlis tersebut diadakan pada waktu pagi hari persekolahan.</p>	<p>tersebut diadakan pada waktu pagi hari persekolahan majlis</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>tersebut diadakan pada waktu pagi hari persekolahan majlis</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: majlis tersebut diadakan pada waktu pagi hari persekolahan</p>
<p>PKN tidak mahu mengganggu lepasan sekolah melanjutkan pelajaran.</p>	<p>tidak mahu mengganggu lepasan sekolah melanjutkan pelajaran PKN</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>PKN tidak mahu mengganggu lepasan sekolah melanjutkan pelajaran</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: PKN tidak mahu mengganggu lepasan sekolah melanjutkan pelajaran</p>
<p>Beribu-ribu ekor burung ruak-ruak menemui sangkar besar itu.</p>	<p>menemui sangkar besar itu beribu-ribu ekor burung ruak-ruak</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>Beribu-ribu ekor burung ruak-ruak menemui sangkar besar itu.</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: beribu-ribu ekor burung ruak-ruak menemui sangkar besar itu</p>

<p>Kawasan tersebut juga menjadi habitat pelbagai jenis burung.</p>	<p>juga menjadi habitat pelbagai jenis burung kawasan tersebut</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p><u>Apakah yang dimaksudkan dengan kawasan tersebut?</u></p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: kawasan tersebut juga menjadi habitat pelbagai jenis burung.</p>
<p>Imran mencuba kemahiran mencantumkan pokok nanas.</p>	<p>mencuba kemahiran mencantumkan pokok nanas Imran</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>mencuba kemahiran mencantumkan pokok nanas imran</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: imran mencuba kemahiran mencantumkan pokok nanas</p>
<p>Agama juga menganggap semua orang islam itu bersaudara.</p>	<p>juga menganggap semua orang islam itu bersaudara agama</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p><u>Apakah yang dimaksudkan dengan semua orang islam itu bersaudara?</u></p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: agama juga menganggap semua orang islam itu bersaudara.</p>
<p>Saya amat bersimpati terhadap nasib ketiga-tiga beranak tersebut.</p>	<p>amat bersimpati terhadap nasib ketiga-tiga beranak tersebut saya</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p><u>Apakah yang dimaksudkan dengan ketiga-tiga beranak tersebut?</u></p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: saya amat bersimpati terhadap nasib ketiga-tiga beranak tersebut.</p>
<p>Bayaran masuk pertandingan boleh dibuat secara tunai sahaja.</p>	<p>boleh dibuat secara tunai sahaja bayaran masuk pertandingan</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p><u>Apakah yang dimaksudkan dengan bayaran masuk pertandingan?</u></p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: bayaran masuk pertandingan boleh dibuat secara tunai sahaja.</p>

Azli telah menerima sijil perkhidmatan cemerlang daripada majikannya.	menerima sijil perkhidmatan cemerlang daripada majikannya Azli telah	<p>Ini adalah sijil yang saya terima sebagai Pegawai <u>Jawatankuasa Perkhidmatan Cemerlang (JPKC)</u> Majlis Perundangan</p> <p>Misal, apat yang dirasmikan salah dari segi rumus bahasa ayat B1.</p> <p>Contoh ayat yang boleh digunakan adalah sebagai berikut: "Saya telah menerima sijil perkhidmatan cemerlang daripada majikannya."</p>
Biotek juga menghasilkan haiwan spesies baharu.	juga menghasilkan haiwan spesies baharu biotek	<p>Ini adalah sijil yang saya terima sebagai Pegawai <u>Jawatankuasa Perkhidmatan Cemerlang (JPKC)</u> Majlis Perundangan</p> <p>Misal, apat yang dirasmikan salah dari segi rumus bahasa ayat B1.</p> <p>Contoh ayat yang boleh digunakan adalah sebagai berikut: "Biotek juga menghasilkan haiwan spesies baharu."</p>
Saya boleh membawa mereka berurusan dengan kerajaan negeri.	membawa mereka berurusan dengan kerajaan negeri saya boleh	<p>Ini adalah sijil yang saya terima sebagai Pegawai <u>Jawatankuasa Perkhidmatan Cemerlang (JPKC)</u> Majlis Perundangan</p> <p>Misal, apat yang dirasmikan salah dari segi rumus bahasa ayat B1.</p> <p>Contoh ayat yang boleh digunakan adalah sebagai berikut: "Saya boleh membawa mereka berurusan dengan kerajaan negeri."</p>
Kita boleh memperingati hari kemerdekaan dengan pelbagai cara.	memperingati hari kemerdekaan dengan pelbagai cara kita boleh	<p>Ini adalah sijil yang saya terima sebagai Pegawai <u>Jawatankuasa Perkhidmatan Cemerlang (JPKC)</u> Majlis Perundangan</p> <p>Misal, apat yang dirasmikan salah dari segi rumus bahasa ayat B1.</p> <p>Contoh ayat yang boleh digunakan adalah sebagai berikut: "Kita boleh memperingati hari kemerdekaan dengan pelbagai cara."</p>
Saya berpendapat kerjaya kepolisan sangat sesuai dengan jiwa saya.	berpendapat kerjaya kepolisan sangat sesuai dengan jiwa saya	<p>Ini adalah sijil yang saya terima sebagai Pegawai <u>Jawatankuasa Perkhidmatan Cemerlang (JPKC)</u> Majlis Perundangan</p> <p>Misal, apat yang dirasmikan salah dari segi rumus bahasa ayat B1.</p> <p>Contoh ayat yang boleh digunakan adalah sebagai berikut: "Saya berpendapat kerjaya kepolisan sangat sesuai dengan jiwa saya."</p>
Pelajar Islam mengambil kesempatan bersolat zuhur di masjid itu.	mengambil kesempatan bersolat zuhur di masjid itu pelajar Islam	<p>Ini adalah sijil yang saya terima sebagai Pegawai <u>Jawatankuasa Perkhidmatan Cemerlang (JPKC)</u> Majlis Perundangan</p> <p>Misal, apat yang dirasmikan salah dari segi rumus bahasa ayat B1.</p> <p>Contoh ayat yang boleh digunakan adalah sebagai berikut: "Pelajar Islam mengambil kesempatan bersolat zuhur di masjid itu."</p>

Ujian ini dijangka akan dikendalikan oleh kumpulan penilai bertauliah.	dijangka akan dikendalikan oleh kumpulan penilai bertauliah ujian ini	<p>Sila masukkan ayat yang tidak malar ini ke dalam ruang yang disediakan dalam bentuk yang betul.</p> <p>Putih Penghijauan</p> <p>Maksud, ayat yang dimasukkan adalah dari segi rumus binomial ayat BBL.</p> <p>Carilah ayat yang betul digunakan ujian ini di antara akan diberikan dalam bentuk yang betul yang akan bertauliah.</p>
Sarawak juga merupakan destinasi pelancongan utama di negara kita	juga merupakan destinasi pelancongan utama di negara kita Sarawak	<p>Sila masukkan ayat yang tidak malar ini ke dalam ruang yang disediakan dalam bentuk yang betul.</p> <p>Putih Penghijauan</p> <p>Maksud, ayat yang dimasukkan adalah dari segi rumus binomial ayat BBL.</p> <p>Carilah ayat yang betul digunakan ujian ini di antara akan diberikan dalam bentuk yang betul yang akan bertauliah.</p>
Mercu gunung api merupakan satu daripada ciri terkemuka taman ini.	merupakan satu daripada ciri terkemuka taman ini mercu gunung api	<p>Sila masukkan ayat yang tidak malar ini ke dalam ruang yang disediakan dalam bentuk yang betul.</p> <p>Putih Penghijauan</p> <p>Maksud, ayat yang dimasukkan adalah dari segi rumus binomial ayat BBL.</p> <p>Carilah ayat yang betul digunakan ujian ini di antara akan diberikan dalam bentuk yang betul yang akan bertauliah.</p>
Langkah ini memupuk nilai murni di kalangan belia.	memupuk nilai murni di kalangan belia langkah ini	<p>Sila masukkan ayat yang tidak malar ini ke dalam ruang yang disediakan dalam bentuk yang betul.</p> <p>Putih Penghijauan</p> <p>Maksud, ayat yang dimasukkan adalah dari segi rumus binomial ayat BBL.</p> <p>Carilah ayat yang betul digunakan ujian ini di antara akan diberikan dalam bentuk yang betul yang akan bertauliah.</p>
Pelantar menggerudi petroleum didirikan di pantai laut china selatan.	mengerudi petroleum didirikan di pantai laut china selatan pelantar	<p>Sila masukkan ayat yang tidak malar ini ke dalam ruang yang disediakan dalam bentuk yang betul.</p> <p>Putih Penghijauan</p> <p>Maksud, ayat yang dimasukkan adalah dari segi rumus binomial ayat BBL.</p> <p>Carilah ayat yang betul digunakan ujian ini di antara akan diberikan dalam bentuk yang betul yang akan bertauliah.</p>
Usaha sebegini dapat menarik minat mereka	dapat menarik minat mereka usaha sebegini	<p>Sila masukkan ayat yang tidak malar ini ke dalam ruang yang disediakan dalam bentuk yang betul.</p> <p>Putih Penghijauan</p> <p>Maksud, ayat yang dimasukkan adalah dari segi rumus binomial ayat BBL.</p> <p>Carilah ayat yang betul digunakan ujian ini di antara akan diberikan dalam bentuk yang betul yang akan bertauliah.</p>

<p>Tepukan diberi pada setiap kali peserta membuat persembahan.</p>	<p>diberi pada setiap kali peserta membuat persembahan tepukan</p>	<p>Sila masukkan ayat yang boleh melebihi 14 perkataan</p> <p><input type="text" value="Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM."/></p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: tepukan diberikan kepada setiap kali peserta membuat persembahan</p>
<p>Kita boleh mendapat maklumat dengan cara menyoal sesuatu pihak.</p>	<p>boleh mendapat maklumat dengan cara menyoal sesuatu pihak kita</p>	<p>Sila masukkan ayat yang boleh melebihi 14 perkataan</p> <p><input type="text" value="Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM."/></p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: kita boleh mendapatkan maklumat dengan cara menyoal pihak kita</p>
<p>Midun menunjuk ke arah depan.</p>	<p>menunjuk ke arah depan Midun</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p><input type="text" value="menunjuk ke arah depan midun"/></p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: midun menunjuk ke arah depan</p>
<p>Anda diajak oleh rakan anda menyaksikan perlumbaan Formula 1.</p>	<p>diajak oleh rakan anda menyaksikan perlumbaan Formula 1 anda</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p><input type="text" value="Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM."/></p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: anda diajak oleh rakan anda menyaksikan perlumbaan Formula 1</p>
<p>Mereka menggunakan pendatang tanpa izin untuk mengaut keuntungan.</p>	<p>menggunakan pendatang tanpa izin untuk mengaut keuntungan mereka</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p><input type="text" value="Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM."/></p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: mereka menggunakan pendatang tanpa izin untuk mengaut keuntungan</p>

<p>Saya harap anda tidak bermasalah untuk bergaul dengan rakan-rakan.</p>	<p>harap anda tidak bermasalah untuk bergaul dengan rakan-rakan saya</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>harap anda tidak bermasalah untuk bergaul dengan rakan-rakan say:</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: saya harap anda tidak bermasalah untuk bergaul dengan rakan-rakan</p>
<p>Anda tidak menghadapi sebarang penyakit.</p>	<p>tidak menghadapi sebarang penyakit anda</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>tidak menghadapi sebarang penyakit anda</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: anda tidak menghadapi sebarang penyakit</p>
<p>Makanan ini menjadi kegemaran pada setiap kali hari raya.</p>	<p>menjadi kegemaran pada setiap kali hari raya makanan ini</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>makanan ini menjadi kegemaran pada setiap kali hari raya</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: makanan ini menjadi kegemaran pada setiap kali hari raya</p>
<p>Beberapa langkah telah dilaksanakan kerajaan untuk memperkukuh integrasi nasional.</p>	<p>telah dilaksanakan kerajaan untuk memperkukuh integrasi nasional beberapa langkah</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>Beberapa langkah telah dilaksanakan kerajaan untuk memperkukuh integrasi nasional.</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: Beberapa langkah telah dilaksanakan kerajaan untuk memperkukuh integrasi nasional.</p>
<p>Rakan anda tidak mahu menggunakan urus niaga melalui perkhidmatan e-dagang.</p>	<p>tidak mahu menggunakan urus niaga melalui perkhidmatan e-dagang rakan anda</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>rakan anda tidak mahu menggunakan urus niaga melalui perkhidmatan e-dagang</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: rakan anda tidak mahu menggunakan urus niaga melalui perkhidmatan e-dagang</p>

<p>Pengerusi meminta semua ahli hadir bagi menjayakan projek tersebut.</p>	<p>meminta semua ahli hadir bagi menjayakan projek tersebut pengerusi</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: meminta semua ahli hadir bagi menjayakan projek tersebut</p>
<p>Setiausaha telah mengedarkan kertas kerja.</p>	<p>mengedarkan kertas kerja setiausaha telah</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>mengedarkan kertas kerja setiausaha telah</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: setiausaha telah mengedarkan kertas kerja</p>
<p>Beliau mengadakan program hari kesihatan.</p>	<p>mengadakan program hari kesihatan beliau</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>mengedarkan kertas kerja setiausaha telah</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: setiausaha telah mengedarkan kertas kerja</p>
<p>Pertubuhan ini juga dapat membentuk belia mengikut acuan kita sendiri.</p>	<p>juga dapat membentuk belia mengikut acuan kita sendiri pertubuhan ini</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: juga dapat membentuk belia mengikut acuan kita sendiri</p>
<p>Sastera tradisional juga boleh dipermudah untuk bacaan murid sekolah rendah</p>	<p>juga boleh dipermudah untuk bacaan murid sekolah rendah sastera tradisional</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: sastera tradisional juga boleh dipermudah untuk bacaan murid sekolah rendah</p>

Para pengunjung bukan sahaja dapat melihat keunikan seni reka menara ini	bukan sahaja dapat melihat keunikan seni reka menara ini para pengunjung	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>Para pengunjung bukan sahaja dapat melihat keunikan seni reka menara ini para pengunjung</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: Para pengunjung bukan sahaja dapat melihat keunikan seni reka menara ini</p>
Kekurangan vitamin A boleh menyebabkan seseorang itu mengalami rabun malam.	boleh menyebabkan seseorang itu mengalami rabun malam kekurangan vitamin A malam	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>Kekurangan vitamin A boleh menyebabkan seseorang itu mengalami rabun malam kekurangan vitamin A malam</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: Kekurangan vitamin A boleh menyebabkan seseorang itu mengalami rabun malam</p>
Ruang ini boleh diturunkan untuk memisahkannya daripada tempat duduk penonton.	boleh diturunkan untuk memisahkannya daripada tempat duduk penonton ruang ini	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>Ruang ini boleh diturunkan untuk memisahkannya daripada tempat duduk penonton ruang ini</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: Ruang ini boleh diturunkan untuk memisahkannya daripada tempat duduk penonton</p>
Kejayaan polis itu melegakan orang ramai	melegakan kejayaan polis itu orang ramai	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>Kejayaan polis itu melegakan orang ramai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: orang ramai melegakan kejayaan polis itu</p>
Keahlian kelab ini terhad	terhad keahlian kelab ini	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>Keahlian kelab ini terhad</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: keahlian kelab ini terhad</p>

Cucu-cucu tidak mahu beredar	tidak mahu beredar cucu-cucu	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>Tidak mahu beredar cucu-cucu</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: cucu-cucu tidak mahu beredar</p>
Semuanya tidak akan tersangkut	tersangkut semuanya tidak akan	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>Tersangkut semuanya tidak akan</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: semuanya tidak akan tersangkut</p>
Aku tidak bersekolah hari ini	tidak bersekolah hari ini aku	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>tidak bersekolah hari ini aku</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: aku tidak bersekolah hari ini</p>
Singapura tidak menganggotai majlis ini	tidak menganggotai majlis ini singapura	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>tidak menganggotai majlis ini singapura</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: singapura tidak menganggotai majlis ini</p>
Bapa tidak pergi ke pekan pagi ini	ke pekan pagi ini bapa tidak pergi	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>ke pekan pagi ini bapa tidak pergi</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: bapa tidak pergi ke pekan pagi ini</p>

Anda tidak bersependapat dengan rakan anda itu	dengan rakan anda itu anda tidak bersependapat	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p><input type="text" value="dengan rakan anda itu anda tidak bersependapat"/></p> <p><input type="button" value="Pohon Penghurai"/></p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: anda itu bersependapat dengan rakan anda itu.</p>
Beliau tidak mengalah	mengalah beliau tidak	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p><input type="text" value="mengalah beliau tidak"/></p> <p><input type="button" value="Pohon Penghurai"/></p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: beliau tidak mengalah</p>
Reka bentuk masjid ini amat menarik	amat menarik reka bentuk masjid ini	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p><input type="text" value="amat menarik reka bentuk masjid ini"/></p> <p><input type="button" value="Pohon Penghurai"/></p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: masjid ini amat menarik reka bentuk</p>
Mereka tidak tahu berbahasa mandarin	berbahasa mandarin mereka tidak tahu	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p><input type="text" value="berbahasa mandarin mereka tidak tahu"/></p> <p><input type="button" value="Pohon Penghurai"/></p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: mereka tidak tahu berbahasa mandarin</p>
Saya tidak suka membuang masa	tidak suka membuang masa saya	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p><input type="text" value="tidak suka membuang masa saya"/></p> <p><input type="button" value="Pohon Penghurai"/></p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: saya tidak suka membuang masa</p>

Saya tidak tahu	tidak tahu saya	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>tidak tahu saya</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: saya tidak tahu</p>
Midun tidak takut	tidak takut Midun	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>tidak takut midun</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: midun tidak takut</p>
Anda tidak prihatin	tidak prihatin anda	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>tidak prihatin anda</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: anda tidak prihatin</p>
Hamid sungguh rendah hati orangnya	rendah hati orangnya hamid sungguh	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>rendah hati orangnya hamid sungguh</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: hamid sungguh rendah hati orangnya</p>
Ibu sungguh baik	sungguh baik ibu	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>sungguh baik ibu</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: ibu sungguh baik</p>

Klia sungguh istimewa	sungguh istimewa klia	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>sungguh istimewa klia</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: klia sungguh istimewa</p>
Susila amat gembira	amat gembira susila	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>amat gembira susila</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: susila amat gembira</p>
Hilmi sangat gembira	sangat gembira Hilmi	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>sangat gembira hilmi</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: hilmi sangat gembira</p>
Hujan amat lebat	amat lebat hujan	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>amat lebat hujan</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: hujan amat lebat</p>
Dia sungguh bijak	sungguh bijak dia	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>sungguh bijak dia</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: dia sungguh bijak</p>

Markahnya paling tinggi	paling tinggi markahnya	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>paling tinggi markahnya</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: markahnya paling tinggi</p>
Lantainya amat bersih	amat bersih lantainya	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>amat bersih lantainya</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: lantainya amat bersih</p>
Kawasannya amat bersih	amat bersih kawasannya	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>amat bersih kawasannya</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: kawasannya amat bersih</p>
Projek tersebut perlu disiapkan dalam jangka masa dua minggu.	disiapkan dalam jangka masa dua minggu projek tersebut perlu	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>Projek tersebut perlu disiapkan dalam jangka masa dua minggu.</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: projek tersebut perlu disiapkan dalam jangka masa dua minggu.</p>
Anda ingin mencari maklumat untuk dimuatkan dalam majalah sekolah anda.	mencari maklumat untuk dimuatkan dalam majalah sekolah anda anda ingin	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>Anda ingin mencari maklumat untuk dimuatkan dalam majalah sekolah anda.</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: mencari maklumat untuk dimuatkan dalam majalah sekolah anda.</p>

<p>Anda sebagai pencadang utama</p>	<p>utama anda sebagai pencadang</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>Utama anda sebagai pencadang </p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: anda sebagai pencadang utama</p>
<p>Udara di situ bersih</p>	<p>di situ udara bersih</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>di situ bersih udara</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: udara di situ bersih</p>
<p>Aliah daripada keluarga sederhana</p>	<p>keluarga sederhana Aliah daripada</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>keluarga sederhana aliah daripada</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: aliah daripada keluarga sederhana</p>
<p>Gambar di sebelah menunjukkan struktur gigi manusia</p>	<p>menunjukkan struktur gigi manusia gambar di sebelah</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>gambar di sebelah menunjukkan struktur gigi manusia</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: gambar di sebelah menunjukkan struktur gigi manusia</p>
<p>Kubah pada bangunan merupakan ciri seni bina islam</p>	<p>merupakan ciri seni bina Islam pada bangunan</p>	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>merupakan ciri seni bina Islam pada bangunan</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: merupakan ciri seni bina Islam pada bangunan</p>

Saya daripada syarikat Agro Ria	daripada syarikat Agro Ria saya	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>daripada syarikat Agro_Ria saya</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: saya daripada syarikat agro_ria</p>
Anda ke klinik gigi	ke klinik gigi anda	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>ke klinik gigi anda</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: anda ke klinik gigi</p>
Internet sebagai sumber maklumat	sebagai sumber maklumat internet	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>sebagai sumber maklumat internet</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: internet sebagai sumber maklumat</p>
Anda sebagai pengarah planetarium	sebagai pengarah planetarium anda	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>sebagai pengarah planetarium anda</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: anda sebagai pengarah planetarium</p>
Mereka di Johor Bahru sekarang	Di Johor Bahru sekarang mereka	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>di Johor_Bahru sekarang mereka</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: mereka di johor_bahru sekarang</p>

Malaysia sebagai destinasi pelancongan	sebagai destinasi pelancongan Malaysia	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>sebagai destinasi pelancongan malaysia</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: malaysia sebagai destinasi pelancongan</p>
--	--	--

Cadangan uji kaji yang salah

1. Pegawai polis itu baik	baik pegawai polis itu	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>Baik pegawai polis itu</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: polis itu baik pegawai</p>
2. Beberapa orang budak perempuan itu murid di sekolah saya	di sekolah saya beberapa orang budak perempuan itu murid	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>Beberapa orang budak perempuan itu murid di sekolah saya</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: di sekolah saya beberapa orang budak perempuan itu murid</p>
3. Bangunan Sultan Ahmad_Samad merupakan bangunan tinggalan masa lampau	merupakan bangunan tinggalan masa lampau bangunan Sultan Ahmad_Samad	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>Bangunan Sultan Ahmad_Samad merupakan bangunan tinggalan masa lampau</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: bangunan Sultan Ahmad_Samad merupakan bangunan tinggalan masa lampau</p>
4. Serangan wabak SARS menyebabkan seluruh dunia menjadi gempar.	menjadi gempar serangan wabak SARS menyebabkan seluruh dunia	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>Serangan wabak SARS menyebabkan seluruh dunia menjadi gempar</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: serangan wabak SARS menyebabkan seluruh dunia menjadi gempar</p>

5. Pak Ismawan sering terasa Iswan ada di sisinya Pak Ismawan ada di sisinya.	sering terasa Iswan ada di sisinya Pak Ismawan	<p>Ujian masa ujian akhir yang telah berakhir 14 pelaksanaan</p> <p>Penyakit yang menyerang Iswan ketika ujian akhir</p> <p>Iskemia Perikardial</p> <p>Iskemia, penyakit yang menyerang Iswan ketika ujian akhir yang merupakan kelainan otot EKG.</p> <p>Kejang otot yang terjadi disebabkan karena otot yang bergerak dengan cepat di pinggang.</p>
6. Penjaga garisan terlebih dahulu sudah mengangkat bendera kuning.	terlebih dahulu sudah mengangkat bendera kuning penjaga garisan	<p>Ujian masa ujian akhir yang telah berakhir 14 pelaksanaan</p> <p>Penyakit yang menyerang Iswan ketika ujian akhir</p> <p>Iskemia Perikardial</p> <p>Iskemia, penyakit yang menyerang Iswan ketika ujian akhir yang merupakan kelainan otot EKG.</p> <p>Kejang otot yang terjadi disebabkan karena otot yang bergerak dengan cepat di pinggang.</p>
7. Hari gawai menjadi perayaan utama di negeri saya.	menjadi perayaan utama di negeri saya hari gawai	<p>Ujian masa ujian akhir yang telah berakhir 14 pelaksanaan</p> <p>Penyakit yang menyerang Iswan ketika ujian akhir</p> <p>Iskemia Perikardial</p> <p>Iskemia, penyakit yang menyerang Iswan ketika ujian akhir yang merupakan kelainan otot EKG.</p> <p>Kejang otot yang terjadi disebabkan karena otot yang bergerak dengan cepat di pinggang.</p>
8. Syarikat tempat Hadi bekerja menyalahkan Hadi atas kerugian tersebut.	menyalahkan Hadi atas kerugian tersebut syarikat tempat Hadi bekerja	<p>Ujian masa ujian akhir yang telah berakhir 14 pelaksanaan</p> <p>Penyakit yang menyerang Iswan ketika ujian akhir</p> <p>Iskemia Perikardial</p> <p>Iskemia, penyakit yang menyerang Iswan ketika ujian akhir yang merupakan kelainan otot EKG.</p> <p>Kejang otot yang terjadi disebabkan karena otot yang bergerak dengan cepat di pinggang.</p>
9. Mereka tidak mengambil peduli akan nilai sambutan hari kebangsaan.	tidak mengambil peduli akan nilai sambutan hari kebangsaan mereka	<p>Ujian masa ujian akhir yang telah berakhir 14 pelaksanaan</p> <p>Penyakit yang menyerang Iswan ketika ujian akhir</p> <p>Iskemia Perikardial</p> <p>Iskemia, penyakit yang menyerang Iswan ketika ujian akhir yang merupakan kelainan otot EKG.</p> <p>Kejang otot yang terjadi disebabkan karena otot yang bergerak dengan cepat di pinggang.</p>
10. Majlis ceramah kerjaya tersebut akan diadakan tidak lama lagi.	tersebut akan diadakan tidak lama lagi majlis ceramah kerjaya	<p>Ujian masa ujian akhir yang telah berakhir 14 pelaksanaan</p> <p>Penyakit yang menyerang Iswan ketika ujian akhir</p> <p>Iskemia Perikardial</p> <p>Iskemia, penyakit yang menyerang Iswan ketika ujian akhir yang merupakan kelainan otot EKG.</p> <p>Kejang otot yang terjadi disebabkan karena otot yang bergerak dengan cepat di pinggang.</p>

11. Dua orang guru sepenuh masa telah ditugaskan mengajar ahli kelab.	telah ditugaskan mengajar ahli kelab dua orang guru sepenuh masa	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>telah ditugaskan mengajar ahli kelab dua orang guru sepenuh masa</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: dua orang guru telah ditugaskan mengajar ahli kelab sepenuh masa</p>
12. Rakan anda memberikan hujahnya tentang simpanan gas asli di negara kita.	memberikan hujahnya tentang simpanan gas asli di negara kita rakan anda	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>memberikan hujahnya tentang simpanan gas asli di negara kita rakan anda</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: rakan anda memberikan hujahnya tentang simpanan gas asli di negara kita</p>
13. Hujah peguam bela itu berasaskan fakta kukuh.	berasaskan fakta kukuh hujah peguam bela itu	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>berasaskan fakta kukuh hujah peguam bela itu</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: hujah peguam bela itu berasaskan fakta kukuh</p>
14. Pelajar lelaki itu tersenyum.	tersenyum pelajar lelaki itu	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>Tersenyum pelajar lelaki itu</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: lelaki itu tersenyum pelajar</p>
15. Pemain bola itu amat lincih.	amat lincih pemain bola itu	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>amat lincih pemain bola itu</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: bola itu amat lincih pemain</p>
16. Perpaduan kaum amat penting dalam usaha mengekalkan keamanan.	amat penting dalam usaha mengekalkan keamanan perpaduan kaum	<p>Sila masukkan ayat yang tidak melebihi 14 perkataan</p> <p>amat penting dalam usaha mengekalkan keamanan perpaduan kaum</p> <p>Pohon Penghurai</p> <p>Maaf, ayat yang dimasukkan salah dari segi rumus binaan ayat BM.</p> <p>Cadangan ayat yang boleh digunakan: perpaduan kaum amat penting dalam usaha mengekalkan keamanan</p>

Lampiran J

Biodata pakar penilai

Nama	Jawatan	Umur	Tahap Pendidikan	Pengalaman sebagai MD/pakar bidang
Ros Silawati binti Ahmad	Pensyarah Jabatan Pengajian Melayu/Munsiyi Dewan	42	Sarjana	10
Prof. Madya Dr. Wan Amizah binti Wan Mahmud	Pensyarah kanan DS54/ Munsiyi Dewan	52	PhD	19
Siti Salmah binti Sulaiman	Guru kanan/ Munsiyi Dewan	51	-	10
Noor Suraya binti Adnan Sallehudin	Pengarah Urusan/ <i>Editor</i> / Munsiyi Dewan	50	Sarjana	10
Prof. Madya Nazlia binti Omar	Pensyarah kanan bidang NLP/Pengkomputeran linguistik	44	PhD	15
Dr. Sabrina Tiun	Pensyarah kanan bidang NLP/Pengkomputeran linguistik	40	PhD	15

Lampiran K

Senarai penerbitan

- Yusnita, M. N., & Zulikha, J. (2012). Parser with Sentence Correction for Malay Language (BM). *2012 International Conference on Information and Knowledge Management (ICIKM)*. 24-26 July. Kuala Lumpur, Malaysia.
- Yusnita, M. N., & Zulikha, J. (2012b). Comparison of Syntax Tree Visualization: Toward Malay Language (BM) Syntax Tree. *2012 International Conference on Information and Knowledge Management (ICIKM)*. 24-26 July. Kuala Lumpur, Malaysia.
- Yusnita, M. N., & Zulikha, J. (2012c). Malay Declarative Sentence: Visualization and Sentence Correction. *Open Systems (ICOS), 2012 IEEE Conference on Open Systems (ICOS)*, 21-24 Oct. Kuala Lumpur, Malaysia.
- Yusnita, M. N., & Zulikha, J. (2014). Parse Tree Visualization for Malay Sentence (BMTutor). *Advancement in Information Technology International Conference (ADVIT)*. 16-18 December. Bandung, Indonesia.
- Yusnita, M. N., & Zulikha, J. (2014b). BMTutor Research Design: Malay Sentence Parse Tree Visualization. *2014 IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*. 28-30 Nov. Pulau Pinang, Malaysia.
- Yusnita, M. N., & Zulikha, J. (2015). Malay Parse Tree Sentence Visualisation (BMTutor): Components and Model. *Malaysian Technical Universities Conference on Engineering and Technology (Mucet2015)*. 11-13 Oct 2015. Johor Bahru, Malaysia.
- Yusnita, M. N., & Zulikha, J. (2014). Parse Tree Visualization for Malay Sentence (BMTutor). *ARPJ Journal of Engineering and Applied Sciences*. 10(3). 1253-1259.
- Yusnita, M. N., & Zulikha, J. (2015). Malay Parse Tree Sentence Visualisation (BMTutor): Components and Model. *ARPJ Journal of Engineering and Applied Sciences*.

Lampiran L

Senarai anugerah

Tarikh	Penyelidik	Tajuk produk	Pameran	Pingat
17-19 Julai 2012	Yusnita Muhamad Noor Prof. Dr. Zulikha binti Jamaludin	BMTutor: Kenali Ayat Bahasa Melayu	Pameran reka cipta, Penyelidikan dan Inovasi Malaysia 2012 (PRPi12)	Perak
25 Mei 2014	Yusnita Muhamad Noor Prof. Dr. Zulikha binti Jamaludin	BMTutor: Kenali Ayat Bahasa Melayu	ICT reserach and innovation expo	Tempat pertama kategori produk dan inovasi
1-2 Jun 2015	Yusnita Muhamad Noor Prof. Dr. Zulikha binti Jamaludin	Exploring Bahasa Melayu Sentence using BMTutor	Inovatif Research, Invention & Application (I-RIA) 2015	Perak
27-28 Oktober 2015	Yusnita Muhamad Noor Prof. Dr. Zulikha binti Jamaludin	BMTutor: Learn Bahasa Melayu Visually	UTeM Research and Innovation Expo 2015 (UTEMeX2015)	Perak

Lampiran M

Surat pengesahan Munsyi Dewan

