

A Framework for Malay Computational Grammar Formalism based on an Enhanced Pola Grammar

Hassan Mohamed^{a,*}, Nur Aisyah Abdul Fataf^a, Tengku Mohd Tengku Sembok^a

^a Cyber Security and Digital Industrial Revolution Centre, National Defence University of Malaysia, Kuala Lumpur, 57000, Malaysia

Corresponding author: hassan@upnm.edu.my

Abstract— In the era of IR4.0, Natural Language Processing (NLP) is one of the major focuses because text is stored digitally to code the information. Natural language understanding requires a computational grammar for syntax and semantics of the language in question for this information to be manipulated digitally. Many languages around the world have their own computational grammars for processing syntax and semantics. However, when it comes to the Malay language, the researchers have yet to come across a substantial computational grammar that can process Malay syntax and semantics based on a computational theoretical framework that can be applied in systems such as e-commerce. Hence, we intend to propose a formalism framework based on enhanced Pola Grammar with syntactic and semantic features. The objectives of this proposed framework are to create a linguistic computational formalism for the Malay language based on theoretical linguistic; implement templates for Malay words to handle syntax and semantic features in accordance with the enhanced Pola Grammar; and create a Malay Language Parser Algorithm that can be used for digital applications. To accomplish the objectives, the proposed framework will recursively formalise the computational Malay grammar and lexicon using a combination of solid theoretical linguistic foundations such as Dependency Grammar. A Malay parsing algorithm will be developed for the proposed model until the formalised grammar is deemed reliable. The findings of this indigenous Malay parser will help to advance Malay language applications in the digital economy.

Keywords— Malay Language; Malay Parser; Dependency Grammar.

Manuscript received 15 Oct. 2020; revised 29 Jan. 2021; accepted 2 Feb. 2021. Date of publication 17 Feb. 2021. International Journal on Informatics Visualization is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.



I. INTRODUCTION

Across the globe, information is stored in electronic and digital forms to computers and the Internet. Aside from images, colour and sounds, most information is coded in natural languages. Computational grammar for the syntax and semantics of natural languages is necessary for computers to process, manipulate and extract meaning from this information [1]. Recently, NLP researchers have been showing interest in the expansion of digital commerce that made online product search possible. It is a popular and successful way for customers to identify their desired products and shop online. Customers use natural language queries to find the products they wish to purchase, and the information retrieval techniques deliver the results [2].

A modelling framework for consumer intent applies contextualisation embeddings through NLP [3]. The stringent requirements of the digital era have brought forth considerable research problems in the fields of dialogue

systems, spoken natural language processing, human-computer interface, and search and recommender systems [4].

The recent research on this NLP niche by [5] contributed to a dataset of 3,540 natural language queries. Most e-commerce systems use English as the transaction medium. Malay Language ranked fifth globally and used widely in Malaysia, Indonesia, Brunei, and Southern Thai. Therefore, it is of critical importance to provide Malay medium as the transaction language.

[6] collected and tested a set of Malay grammar using the Pola Grammar technique, which this research used to begin the project for syntactic features. [6] grouped a sentence into adjunct, subject, post-subject, conjunction and predicate, and obtained an F-score index of between 73% to 93%. Their work is small-scale research based on the syntactic feature of Malay grammar. The development of the prototype used a “collection of thesis abstracts” as the training corpus, which did not include many other Malay sentence patterns. This research will extend the work by [6] to improve the pola

grammar for the syntactic and semantics features using the progressive test method.

The Pola Grammar approach help make sense of the grammatical relationships of Malay sentences. The functional structure of the subject's semantic role is the relations consisting of lexical mappings. Lexical mapping requires dictionaries that manage lexical information through rule formalisms that analyse syntactic and semantic aspects in a linguistically justified; and yet computationally viable manner to attain universality. These rules allow us to make linguistic generalisations with a precise computational interpretation [7].

The parsing process uses the disambiguation of the part-of-speech for the tokens (lexical mapping) and the holistic framework of Malay syntactic and semantic patterns in various text domains to improve the precision or recall and the F-score index. Therefore, the researchers hope to build an enhanced formalism as the progressive experimental framework to analyse and accumulate the patterns by stages. Since [6] has collected and tested a set of Malay grammar patterns using the Pola Grammar technique, one of the improvements we will adapt is disambiguating the part-of-speech of the tokens using the part-of-speech (POS) tagger tool produced by [8]. This tagger will help identify the 'template of Malay' words in handling syntactic and semantic features.

The basic Malay pola grammar produced by [9] and [10] are as follows:

- Noun Phrase + Noun Phrase
- Noun Phrase + Verb Phrase
- Noun Phrase + Preposition Phrase
- Noun Phrase + Adjective Phrase

[6] extended the pola grammar for computational form into the following five polas.

- i. adjunct
- ii. subject
- iii. post-subject
- iv. conjunction
- v. predicate

Handling numbers, tenses, gender and cases present challenging problems in parsing Malay sentences because these features are not as explicit as in English or Arabic. The parsing process involves many inferences, and one of the major problems in parsing a document is handling anaphora. Hence it is essential to have a sound framework for the computational handling of Malay words, phrases, sentences and essays. The theoretical foundations in this research are Phrase Structure Grammar formalism [11-12], Logical Categorical Grammar [13], a logical deduction with first-order predicate logic [14-15] and languages such as Prolog as the implementation platform. Phrase Structure Grammar is very much suited to the Malay language phrase and sentence structures. Logical Categorical Grammar can handle syntax and semantics simultaneously with well-defined words and grammar templates using Lambda Calculus. The development of Malay Language grammar in a computational framework is an underexplored territory.

This framework will enhance the pola grammar with syntactic and semantic features using the progressive test for documents such as the Parliament Malaysia Hansard,

newspapers and textbooks to produce a new theoretical foundation for Malay Computational Grammar. Other languages, especially English, have applied Phrase Structure Grammar formalism [11-12] and Dependency Grammar Formalism [16].

II. MATERIALS AND METHOD

We formulated two hypotheses in developing this framework. First, by using the enhanced pola grammar with syntactic and semantic features, a new theoretical framework for Malay Computational Grammar will be established. Second, by incorporating the semantic features, the formalism can cater both syntactic and semantic analysis synchronously. We hope that implementing these hypotheses will improve the effectiveness of Malay based applications such as sentiment analysis, speech for robot etc.

The proposed framework will bridge at least three knowledge gaps. The first is building appropriate theoretical linguistic foundations for the Malay language for use in the computational framework for digital era applications. Filling this gap will contribute to building a linguistic computational framework for the Malay language based on a fusion of sound theoretical linguistic foundations such as Phrase Structure Grammar, Dependency Grammar and Lambda Calculus formalism into enhanced pola grammar with syntactic and semantic features. This formalism is because the Malay language relies heavily on a pragmatic level of inference which is loose in structure [17], reflecting the indeterministic nature of word class transitions in a sentence. This problem creates difficulties in determining explicit syntactic rules for the Malay language.

The second knowledge gap concerns using effective Malay words as a 'template' to determine syntactic and semantic features. Bridging this gap will help verify the templates' effectiveness for Malay words and grammar in handling the language's syntax and semantic features based on the enhanced pola grammar. For example, *penjodoh bilangan*, or count numeration in Malay, has a unique meaning when used as an object in a sentence and functions like any other noun. However, when used as classifiers, these words count things. Occasionally, words that serve as classifiers retain their original meaning as a noun, which is traceable, but this is not always the case.

The third knowledge gap concerns designing algorithms for developing a Malay language parser appropriate for various applications in the digital era, such as retrieval systems, Q&A systems and knowledge-based systems. Bridging this gap will contribute to designing a Malay language parser algorithm for various applications in the digital age, such as e-commerce and sentiment analysis.

A. The Status of Current Malay Computational Grammar Research

The primary research and development work on computational Malay Language processing is the Example-Based Machine Translation (EBMT) project carried out by USM and MIMOS (utmk.cs.usm.my:8080/ebmt-controller/index.jsp). This project was an Example-based Machine Translation between Malay and English sentences with the bilingual knowledge bank constructed from a Malay-English bilingual corpus. Each pair of the English-Malay translations was codified into Synchronous Structured String Tree Correspondence (S-SSTC). Therefore, it is only appropriate for translation. The approach did not emphasise the morphology, syntactic and semantics of words and Malay grammar.

[8] conducted a shallow-level sentence passing for Malay language, POS tagging, using the Hidden Markov Model with Malay morphological features to assign Malay words with their part of speech tag. This work focused on tagging unknown words and relied on a morphological level without going deeper into the syntactic level of a sentence. They resolved ambiguous words using a statistical technique. The research on POS tagging by [18] relied on Malay morphology to develop a morphological inference analyser based on rules. [19] investigated the performance of the supervised ML approaches in tagging Malay words and the effectiveness of the affixes-based feature patterns. [20] developed a Malay POS tagger that relied on foreign languages (such as English), which did not look at the original Malay linguistic features. The latest research on Malay POS tagging and Named Entity Recognition (NER) was by [21] to develop POS tags and rule-based extraction for the narrator's name in the Malay Hadith texts. This research sought to determine the relationship between hadith narrators to construct a narrator's chain. [22] proposed a direct alignment framework for aligning English and Malay news documents.

[23] researched a Malay parser for Malay language teachers to use to help students having difficulties comprehending the grammar and phrase structure. This research used simple sentences from *Munsyi Dewan* to design the inference rules. The research had a specific objective and did not cover broad application domains. The latest development for a Malay parser was in Indonesian Malay [25], which used Probabilistic Context-Free Grammar (PCFG). However, Indonesian Malay is different from Malaysian Malay on lexical and, sometimes, syntactic and semantic levels. Other than Malay parser, there was a text summarization work by [25] which focused on discovering human compression pattern from the developed Malay summary corpus to improve readability and informativeness of the produced summary. [26] determined the Malay anaphoric *nya* (it), and proposed algorithms and Malay Anaphoric Resolution architecture.

All existing works on Malay Computational Linguistics are small-scale experimental projects. Therefore, there is no advancement in research or application development for the Malay language, such as E-commerce applications, Malay Sentiment Analysis, Knowledge Discovery, Automatic Question Answering Systems, Information Retrieval, and Meaning-based Search Engines.

B. Background Theory and Applied Formalism

The grammatical description of a natural language in a computational environment is known as computational grammar. A very high-level programming language like Prolog could serve as a model for the formalism needed for grammar development or grammar engineering, which is a challenging process. More than 250 million people in the Malay Archipelago speak the Malay language, which has complex morphological, syntactic, and semantic aspects that must be addressed using a computational framework. While semantics pertains to meaning, the syntax is a part of grammar. The syntax rules ensure the grammatical accuracy of a sentence; the semantics rules determine how lexicons, grammatical structures, and other sentence components work together to convey meaning. The formalism of the Malay language is not as easy to understand as in other languages, such as English, which have different grammar, lexicons, and grammatical structures.

Natural language refers to all languages spoken by human beings throughout the globe. [27] believe that "... the human language, the first and the foremost product of the human mind...". Wittgenstein, who is a philosopher, introduced the picture theory of meaning to emphasise that language represents the reality of what one perceives in their mind [28]. The scientists and philosophers pointed out the significance of language in human lives for their intellect, cognition and social engagement.

Before the NLP technique was used in the application domains, such as e-commerce, other languages already had computational grammar to process syntax and semantics. Therefore, NLP tools, such as the parsers, POS taggers and NER of these languages, can be a basis for developing domain applications for the present-day digital age.

One of the early formalisms is the advanced Definite Clause Grammar (DCG) formalism, which we used or modified to develop the framework for this project before considering Dependency Grammar. DCG has a smooth and seamless implementation in the environment of the declarative very high-level language such as Prolog, which facilitates its implementation. Dependency Grammar is similar to the semantic relation between words in a sentence and is appropriate for computational semantic analysis. We believe that the research output for this framework will contribute to employing the Malay language in the digital economy, and the indigenous parser can be applied in many application domains. Further, the output could be a tool for studying the Malay language in the digital world.

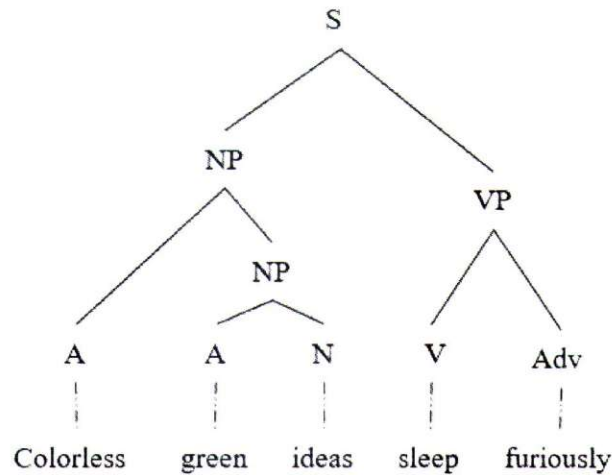


Fig. 1 Phrase tree structure

1) Phrase Structure Grammar

Phrase structure grammar is a form of rewrite rule for explaining the syntax of a particular language [9]. The term rewriting rule in computer science refers to a broad range of techniques for substituting new terms for the subterms in a formula. By dissecting a sentence in natural language into its lexical and phrasal components, we can use phrase structure to understand the overall sentence structure. The grammar rules for phrase construction typically take the following form.

$A \rightarrow B C$ means that constituent A comprises sub-constituents B and C.

Some examples for English are as follows,

$S \rightarrow NP VP$

$S \rightarrow NP VP$

$NP \rightarrow (Det) N1$

$N1 \rightarrow (AP) N1 (PP)$

(S = Sentence; NP = Noun Phrase; Det = Determiner; N1 = Noun Terminal (word); PP = Prepositional Phrase)

Many correct English sentences can be produced by using the symbol S to start a sentence, applying the grammar rules sequentially, and using replacement rules to replace abstract symbols with actual words. Any statement created in this manner is regarded syntactically accurate but meaningless, as in the well-known example "Colorless green ideas sleep furiously." This sentence shows how phrase structure grammar rules can produce a sentence correctly even though the sentence lacks semantic meaning or is erroneous. The sentence can be dissected into its component pieces by applying the above phrase construction rules, as shown in Figure 1.

2) Dependency Grammar

Dependency grammar (DG) is a grammatical theory first articulated by [31]; it is based on the dependency connection [30] instead of the constituency relation of phrase structure. Dependency is the idea that words are linked to one another in a specific way, and the verb is the focal point of the phrase.

Regarding the directed links, every other word is directly or indirectly related to the verb (the centre).

A dependency structure's relationship between a word (the head) and its dependents is what defines it. Due in part to the absence of a finite verb phrase constituent, dependency structures are flatter than phrase structures and thus ideally suited for analysing languages with unconstrained word order, such as Warlpiri or Czech. DGs use a variety of conventions to represent dependencies. Figure 2 shows some of the DG schemata conventions.

Predicates and their arguments help us understand semantic dependencies [32]. A predicate's arguments are subject to it in terms of semantics. Figure 3 shows the word order in the following examples, which demonstrates the typical syntactic dependencies, and the arrows denote the semantic dependencies. In tree (a), arguments John and Mary rely syntactically on the predicate "likes" and depend on it. This dependency indicates an overlap and directionality between the syntactic and semantic dependencies (down the tree). However, attributional adjectives are predicates that take their head noun as their argument. Therefore, in tree (b), "small" is a predicate that takes "ball" as its one argument; the semantic dependency points up the tree and thus contradicts the syntactic dependency.

The conditions in (c) are similar, where the prepositional predicate "on" accepts two arguments, "cup" and "table". One of these semantic dependencies points down the syntactic hierarchy, while the other does the opposite. Finally, the predicate "to take" in (d) accepts only one argument, "Backer". However, it is not directly related to "Backer" in the syntactic hierarchy, proving that semantic dependency is completely separate from syntactic dependencies.

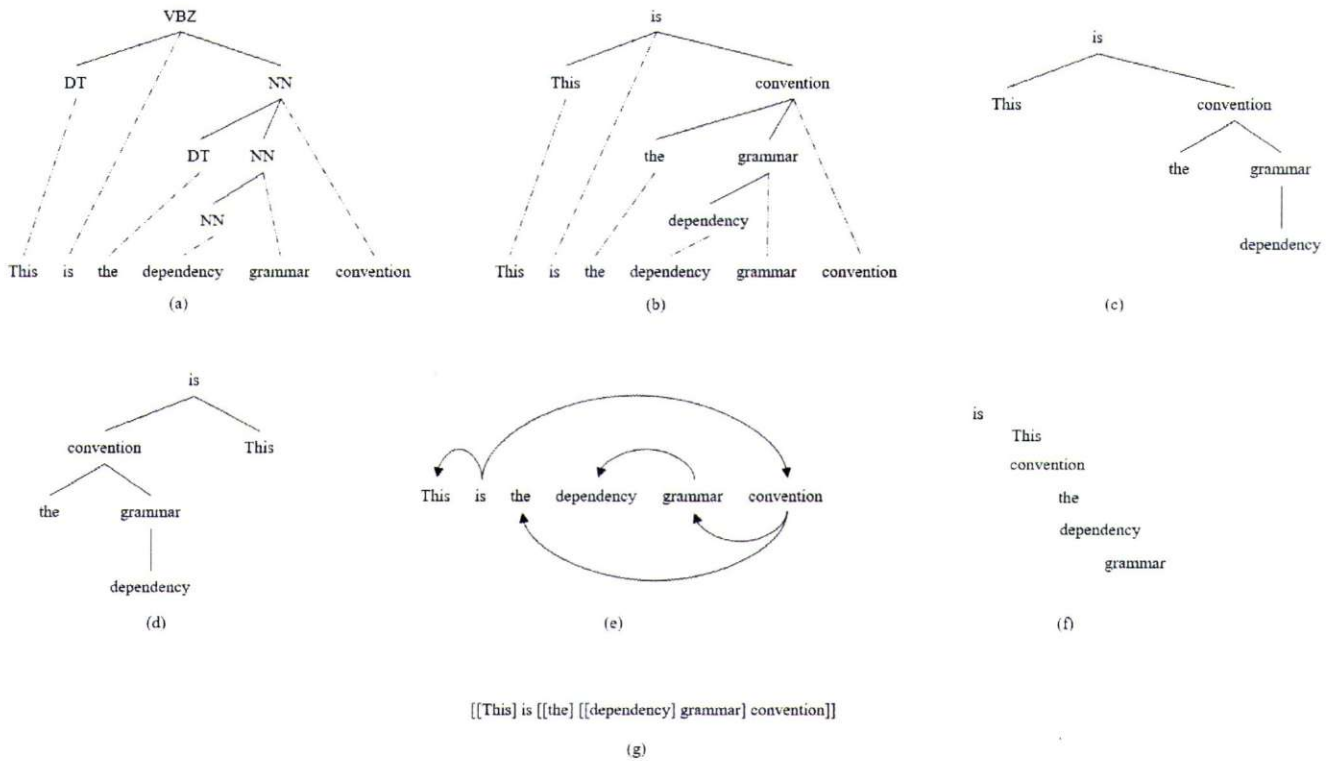


Fig. 2 Dependency grammar structure conventions

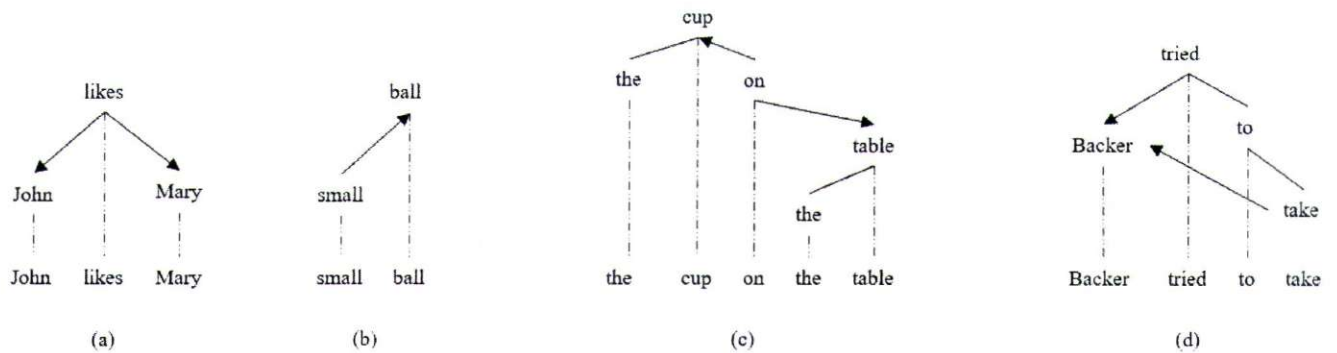


Fig. 3 Semantic Dependency

3) Lambda Calculus

Logical Categorical Grammar can handle syntax and semantics simultaneously by using a well-defined template of words and grammar based on Lambda Calculus, a theory of functions as rules proposed by Church in 1930 [33]. Lambda calculus is well-known as the world's smallest programming language expressing concepts as mathematical functions that may represent any computable issue [34]. It is a valuable tool for defining and verifying programme properties [35]. All search engines must have a methodology for comprehending the content and semantics of a query, and lambda calculus is one potential method for describing semantic representation [36]. This research used Lambda calculus to describe the semantic representations for lexical items and the translations of such expressions in terms of substructure translation in syntax. The researchers believe that the development of

Malay language grammar in a computer framework using lambda calculus is an underexplored territory.

4) Grammar Development Approach

The best progressive test for improving Pola Grammar is based on a combination of deductive and inductive methods. The inductive method relies on documents, and the deductive is based on theoretical linguistics. The variety of documents used in the development and test increase the possibility of obtaining more grammatical patterns complying with and contributing to novel theoretical linguistics suitable for various applications in the digital age [37].

III. RESULT AND DISCUSSION

This section presents a detail discussed of the proposed framework. The proposed formalism and linguistic research framework comprise five phases, as shown in Figure 4.

Phase 1: Thorough Comparative Analysis of the existing resources.

- Natural language processing and understanding based on the theoretical foundation of other languages.
- Carry out the work on Malay Language computational linguistics.
- Parsing algorithms in Natural Language Processing.
- Application of descriptive languages, such as Prolog, CLIPS and NLTK, in Natural Language Processing

Phase 2: Formalisation of problems

The formalisation of problems in developing computational grammar and lexicon for the Malay language requires the researchers and Malay grammar experts to have discussions in a workshop to achieve in-depth insights.

Phase 3: Formalisation of solutions

The objectives for the formalisation of solutions are as follows.

- Model a computational Malay grammar and lexicon using a fusion of sound theoretical linguistic foundations, such as Phrase Structure Grammar, Dependency Grammar and Lambda Calculus formalism.
- Develop the templates for Malay words and grammar to handle the syntax and semantics features of the language using the proposed framework.
- Parsing algorithm for the model proposed in the platform.

Phase 4: Implementation and proof of concept

- Develop the prototype and proof of concept.
- Develop the Malay computational grammar using the formulated templates.
- Collect and populate the Malay words using the formulated templates.

Phase 5: Evaluate the prototype

- Establish formal proof.
- Establish experimental proof based on the collected data.

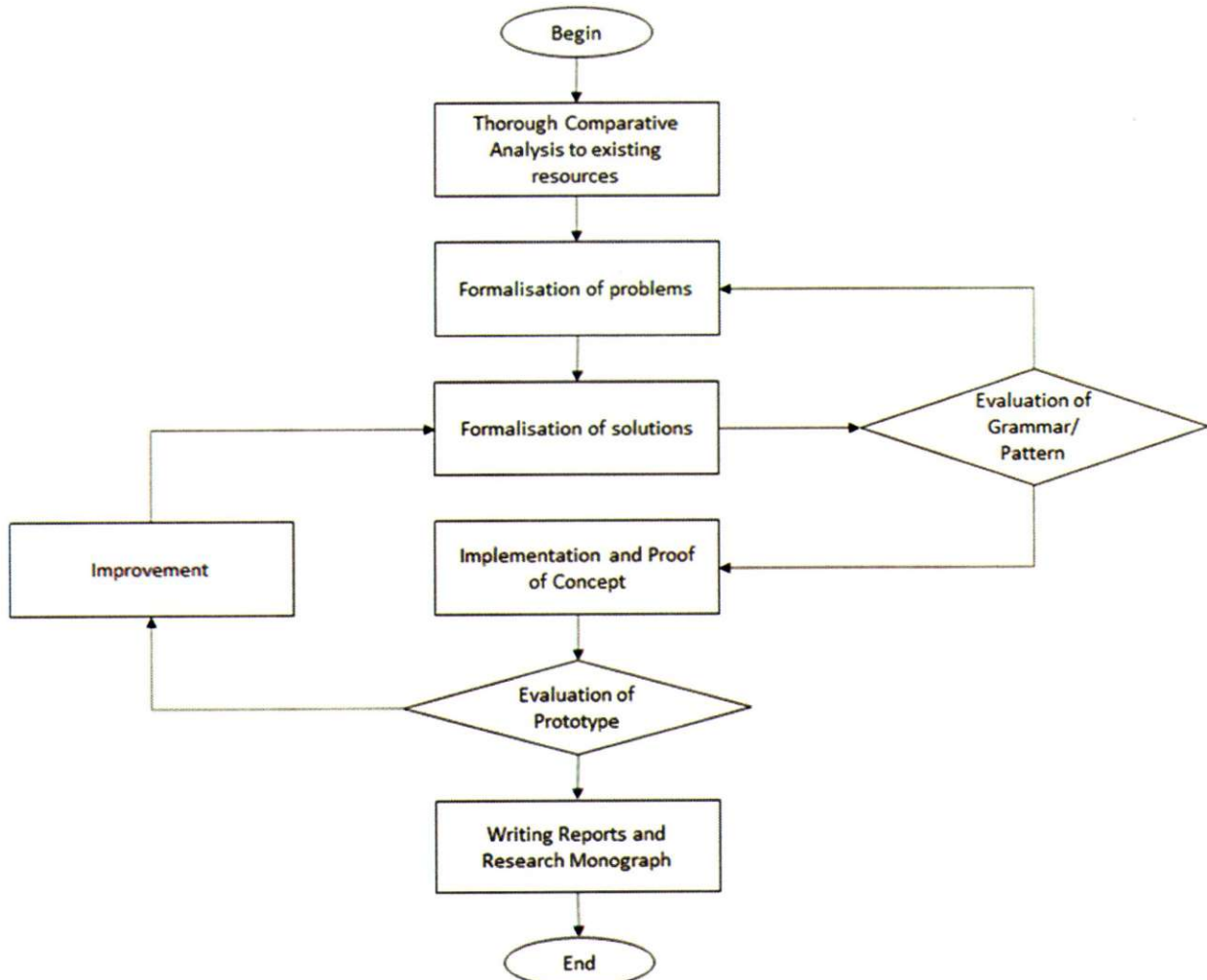


Fig. 4 Flow chart of the proposed formalism framework

IV. CONCLUSION

This paper proposed a formalism framework based on the enhanced Pola Grammar with syntactic and semantic features for the Malay language. This objective was formulated following the theoretical linguistic foundations of the Malay language to implement the templates for Malay words to handle the syntax and semantic features complying with the enhanced Pola Grammar. Another objective was to develop a Malay Language Parser Algorithm appropriate for various applications in the digital age. Therefore, the researchers recursively formalised the computational Malay grammar and lexicon based on a fusion of sound theoretical linguistic foundations, such as Dependency Grammar, to achieve the objectives. The research outcomes contribute to the applications of the Malay language in the digital economy with the manifestation of an indigenous Malay parser.

ACKNOWLEDGEMENT

We thank the National Defence University of Malaysia (UPNM) for supporting the publication of this paper through the Center for Research and Innovation Management (PPPI) and funding the project through the research grant UPNM/2022/GPJP/ICT/1.

REFERENCES

- [1] K. Dashtipour, M. Gogate, J. Li, F. Jiang, B. Kong, A. Hussain. A hybrid Persian sentiment analysis framework: Integrating dependency grammar based rules and deep neural networks, 2020. 380: 1-10
- [2] R.R. Iyer, R. Kohli, S. Prabhumoye. Modeling Product Search Relevance in e-Commerce. 2020. <https://arxiv.org/abs/2001.04980> - [24 March 2021]
- [3] H. Elhabbak, B. Descamps, E. Fischer, S. Athanasiadis, S. Contextualisation of eCommerce Users. 2020. <https://arxiv.org/abs/2011.01874> - [24 March 2021]
- [4] E. Agichtein, D. Hakkani-Tür. ConVERSE'20: The WSDM 2020 Workshop on Conversational Systems for E-Commerce Recommendations and Search. WSDM '20, Houston, TX, USA, pp 897 – 898. 2020.
- [5] A. Papenmeier, D. Kern, D. Hienert, A. Sliwa, A. Aker, N. Fuhr. Dataset of Natural Language Queries for E-Commerce. CHIIR '21, March 14–19, 2021, Canberra, Australia, pp 307 – 311.
- [6] M. M. M. Juzaidin., A.A.A. Ahmad, F. Ghani. Pola Grammar Technique for Grammatical Relation Extraction in Malay Language. *Malaysian Journal of Computer Science, University Malaya*. 2006.
- [7] B. Li, J. Cheng, Y. Liu, & F. Keller. Dependency Grammar Induction with a Neural Variational Transition-Based Parser. Proceedings of the AAAI Conference on Artificial Intelligence, 2019. 33(01), 6658-6665.
- [8] M. Hassan, O. Nazlia, M.M.M. Juzaidin. The Effectiveness of Using Malay Affixes for Handling Unknown Words in Unsupervised HMM POS Tagger. *International Journal of Engineering & Technology*, 7 (4.29), 9-12, 2018.
- [9] O. Asmah, Morfologi-sintaksis Bahasa Malayu (Malay) dan Bahasa Indonesia: Satu Perbandingan Pola, *Dewan Bahasa dan Pustaka*, Kuala Lumpur, 1968.
- [10] N.K. Safiah, M.O. Farid, M. Hasihm, A.H. Mahmood, A. Tatabahasa Dewan, *Dewan Bahasa dan Pustaka*, Kuala Lumpur. 1993
- [11] J. Zhou and H. Zhao. Head-Driven Phrase Structure Grammar Parsing on Penn Treebank. 2020. Cite as: arXiv:1907.02684 [cs.CL] <https://doi.org/10.48550/arXiv.1907.02684>
- [12] E. Seifossadat and H. Sameti. Stochastic Data-to-Text Generation Using Syntactic Dependency Information. *Computer Speech & Language*. 2022. 76: 101388
- [13] T.M.T. Sembok. A Robust Parsing and Translation Strategy for Categorical Grammar, *Sains Malaysiana*, Jld.19(4), 1990.
- [14] R.A. Kadir, T.M.T. Sembok, B.Z. Halimah. Logical Deduction Inference on Logic Representation in Open-ended Question Answering Process. *Seminar Siswazah 2005 FTSM*. 2005, PERMATA, Bangi, Malaysia. 2005.
- [15] R.A. Kadir, T.M.T. Sembok, B.Z. Halimah. A Logical Deduction to First Order Logic for Use in an Intelligent Reading Comprehension. Proceeding of ROVISP'05. pp 819-823. 2005, Penang, Malaysia.
- [16] B. Li, J. Cheng, Y. Liu & F. Keller. Dependency Grammar Induction with a Neural Variational Transition-Based Parser. Proceedings of the AAAI Conference on Artificial Intelligence. 2019. 33(01), 6658-6665. <https://doi.org/10.1609/aaai.v33i01.33016658>
- [17] G. Knowles, M.D. Zuraida. Tagging a Corpus of Malay Text and Coping with Syntactic Drift. Proceedings of the Corpus Linguistics. Centre for Computer Corpus Research on Language, pp. 422–428. University of Lancaster. 2003.
- [18] B. Ranaivo-Malançon. Malay Lexical Analysis Through Corpus-based Approach. Proceedings of International Conference of Malay Lexicology and Lexicography (PALMA), Kuala Lumpur, Malaysia. 2005. [http://eprints.usm.my/10608/1/Language_Identifier_Converter_For_Malay_and_Indonesian_Languages_\(PPSKomputer_\)_2005.pdf](http://eprints.usm.my/10608/1/Language_Identifier_Converter_For_Malay_and_Indonesian_Languages_(PPSKomputer_)_2005.pdf) [29 Julai 2015].
- [19] G. Shamsan, N.M.Z. ahmad; O. Nazlia; A. Salwani. Part-of-speech (pos) tagger for malay language using naïve bayes and k-nearest neighbor model. *Journal of Critical Reviews*. 2020. 7(16): 248.
- [20] Z. Norshuhani, A. Oxley, A.B. Zainab, S.A. Farhan. A Lazy Man's Way to Part-of-Speech Tagging. D. Richards and B. H. 2012.
- [21] N.A. Rahman, N.K. Ismail, Z.M. Nor, M.N. Alias, M.S. Kamis, N. Alias, Tagging Narrator's Names in Hadith Text. *Journal of Fundamental and Applied Sciences*. 2017. <http://dx.doi.org/10.4314/jfas.v9i5s.21>
- [22] N.A. Nasharuddin, M.T. Abdullah, A. Azman, R.A. Kadir. A framework for English and Malay cross-lingual document alignment method. *International Journal of Advanced Trends in Computer Science and Engineering*. 2019. 8(1.3): 190-195
- [23] Y. Maisarah. MYPARSER: A Malay Text Categorization Toolkit Using Inference Rule. Master Thesis, UTHM. 2013.
- [24] D.E. Cahyani, L. Gumilar, A. Pangestu. Indonesian Parsing using Probabilistic Context-Free Grammar (PCFG) and Viterbi-Cocke Younger Kasami (Viterbi-CYK). 3rd International Seminar on Research of Information Technology and Intelligent Systems (ISRITI). 2020. DOI:10.1109/ISRITI51436.2020.9315395 Corpus ID: 231616132.
- [25] A. Suraya. A Pattern-Growth Sentence Compression Technique for Malay Text Summarizer. PhD Thesis, USM. 2018.
- [26] K.M.N. Huzaimi, S.A.M. Noah, Shahrul, M.J.A. Aziz. Classification of short possessive clitic pronoun nya in malay text to support anaphor candidate determination. *Journal of Information and Communication Technology*. 2020. 19(4): 513 - 532
- [27] K. Popper. All Life is Problem Solving. Translated by Patrick Camiller. Routledge, London and New York. 1999.
- [28] D. Keyt. Wittgenstein's Picture Theory of Language. *The Philosophical Review* 73(4), 493-511. 1964.
- [29] A. Lopopolo, A. van den Bosch, K.M. Petersson, & R.M. Willems. Distinguishing syntactic operations in the brain: Dependency and phrase-structure parsing. *Neurobiology of Language*. 2021. 2(1): 152–175.
- [30] T. Osborne. NPs, not DPs: The NP vs. DP debate in the context of dependency grammar. *Acta Linguistica Academica* 68. 2021. 3: 274–317.
- [31] L. Tesnière, Elements of structural syntax. Translated by Timothy Osborne and Sylvain Kahane. Amsterdam & Philadelphia, PA: John Benjamin. 2015
- [32] I. Mel'čuk. Levels of dependency in linguistic description: Concepts and problems. 2003. In Ágel et al., 170–187.
- [33] A. Church. Review of Haskell B. Curry. The Inconsistency of Certain Formal Logics. *The Journal of Symbolic Logic*. 1942. 7(4): 170-171.
- [34] B. Almeida, A. Mordido, P. Thiemann, V.T. Vasconcelos. Polymorphic lambda calculus with context-free session types. *Information and Computation*. 2022
- [35] E.G. Daylight. The halting problem and security's language-theoretic approach: Praise and criticism from a technical historian. *Computability*. 2020. 10: 1-18.
- [36] H. Singh. Visualizing and Computing Natural Language Expressions: Through a Typed Lambda Calculus λ . In: Bajpai, M.K., Kumar Singh, K., Giakos, G. (eds) Machine Vision and Augmented Intelligence—Theory and Applications. Lecture Notes in Electrical Engineering. 2021. 796. Springer, Singapore.
- [37] M. Nasir, C.I. Ezeife. Semantic enhanced Markov model for sequential E-commerce product recommendation. *Int J Data Sci Anal*. 2022. <https://doi.org/10.1007/s41060-022-00343>