VOL. 10, NO. 3, FEBRUARY 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences


www.arpnjournals.com

A CONCEPTUAL MODEL TO MANAGE LEXICAL AMBIGUITY IN MALAY TEXTUAL REQUIREMENTS

Hazlina Haron[1], Abdul Azim Abdul Ghani[2], Hazliza Haron[2]
[1]Faculty of Computer Science and Information Technology, UPM Selangor, Malaysia
[2]Faculty of Business Administration, UiTM (Perak) Seri Iskandar Perak, Malaysia
Email: hliza457@perak.uitm.edu.my

ABSTRACT

Ambiguity is one of the most crucial problem in natural language. When a sentence can be interpreted in more than one ways by different sets of readers, the impact is huge. It would involve many stages in system development life cycle from requirement elicitation to requirement transformation, system design, system coding as well as the end product. It is very important to ensure requirement requested by the users accurately transformed onto the end system as per desired. More often than not, the understanding between users and system developer is not tele. In this paper, we attempt to publish a conceptual model on managing lexical ambiguity to reduce the possibility of misinterpretation errors in Malay sentences. The approach that we used is by identifying potential Malay vague words based on vague criteria and mapped onto constructed Requirement Ambiguity Attributes (RAA). We designed a Model of Vagueness that has the elements of vague criteria that is mapped with RAA as a hybrid method to detect potential ambiguous sentences.

Key words: linguistic ambiguity, Malay ambiguity, syntactic analysis, vague concept, Malay grammar.

INTRODUCTION

In system requirement specification (RS), linguistic ambiguity is often ignored or mistakenly unacknowledged. This results in misunderstanding from both users and system developer's side, hence contributes to a failed system. The impact is on the extended project timeframe and lead to cost-burst. A high quality of a requirement contributes to a successful, cost-effective creation of software system. RS is usually written in natural language (NL) as it is most flexible and easy to understand across discipline and domain. However, it is also prone to language error. One of the known errors is ambiguity in the sentences. It happens when the statements are not well written in a precise manner. Sentence ambiguity occurs when there is more than one interpretation in a sentence that is being read by different sets of people. Linguistic ambiguity can come from many sources, among others are multiple word senses (Burg, 1989), syntactic and structural ambiguity of sentences (Burg, 1989) such as negations and misused of quantifiers (Erik Kamsties & Paech, 2000), long-ranged relationship in terms of word referencing (Burg, 1989; Grenat & Taher, 2008), imprecise usage of words (Burg, 1989), misconception of word meanings (Berry, Kamsties, Krieger, & Lee, 2003), customers don't really know what they want and communication as well as knowledge gap between customers, software engineers and project managers (Yang et al., 2008).

Our aim is to assist in supporting business users and software developer to better understand NL sentences in Malay textual requirement before a new system can be built or enhanced. This is because RS acts as the medium between the two parties and therefore it is significant for both to have a same understanding and interpretations. We intend to develop an automated prototype tool that provide the facilities to auto-detect potential ambiguity in Malay textual requirements. This tool will assist business analysts or requirement engineers to check on their written requirements to reduce or avoid using ambiguous words that contributes to multiple interpretation and ambiguity.

BACKGROUND AND MOTIVATION

Ambiguity

There are numerous researches that have conducted studies concerning ambiguity. Quite a number of researches agree that manipulating natural language processing (NLP) method is one of the best techniques to combat ambiguity. (Ambriola & Gervasi, 1997) proposed on using NLP technique in recognizing lexicons and semantic ambiguity for the process of transforming textual information into formal requirements. The technique has three main elements; 1. Inputs preparation that involves a library of glossary to produce list of significant terms, classify them, synonyms. 2. Fuzzy matching that uses a user-defined MAS-rule (model, action, substitution) where requirements will be tagged and model matching will be used. 3. Action that translates into text once it matches the rules. The technique enforced on explicitness, categorization and classification of entities and their relationships forced users to express their requirements in clear manner. In addition, due to a well-defined semantics in corresponding to MAS-rule enable a certain incomplete requirements to be spotted and detected easily. Likewise, (Huyck & Abbas, n.d.) developed a prototype of NLP tool to detect ambiguity that uses NLP parser and NLP lexical analysis. It parses the specifications and flags any sentences that have multiple interpretations. It emphasized on syntactical and semantic structure in RS. It based on LINK parser (Lytinen 92) that uses chart parser (Allen 87). The input to the LINK system is a grammar, a lexicon and sentences. LINK produces a chart describing all legal

grammar rule applications over the given sentences. The system was tested against 10 requirements sentences.

On the contrarily, (S. F. Tjong, Hartley, & Berry, 2007) uses different approach in reducing requirements ambiguity. They presented a method of how ambiguity can be avoided or reduced by setting up requirement guiding rules. These rules have to be followed which failing to do so, will affect the clarity of the requirements. Each rule offers an alternative sentence pattern and wordings that is less ambiguous and more precise. This work is somehow has a similar approach with (Nuseibeh, Easterbrook, & Russo, 2001). It uses a set of consistency rules that captures the set of relationship which is abided by a set of requirement descriptions. SLR & PROMELA are used to verify requirements where it translates informal specifications into formal notation. PROMELA also is used to verify existing design to know whether or not the proposed design meets the requirement.

From another perspective, (E Kamsties, 2005) proposed to use a scenario-based reading technique for spotting ambiguity before the formal specification is being developed. This is to avoid too much modification that needed to be done after transforming informal requirements into formal specification. It uses several RS languages (RSL) such as Focus, SCR, OCTOPUS, ROOM, Statemate, UML, informal RS and CASE tool. However, the evaluation of the research was done based on students' perspective. Students are those not the real users of a RS and their ability in analyzing and developing requirement model might be limited and can be questionable. It is advisable that author evaluate his method using real users.

(Chantree, Nuseibeh, Roeck, & Willis, 2005) tested their hypothesis on the occurrence of distributional similarities that is likely to be coordinated before modifiers takes scope. Consider a sentence *"old boots and shoes"* as an example. The words *'boots'* and *'shoes'* are perceived first before modifier *'old'* being applied. Distributional similarities are when two words falls under same categories and have similar meaning. The author then enhanced their research by identifying and distinguished nocuous and innocuous ambiguity where the methods that are being used are human judgement, application based as well as heuristic analysis. (Al-Fawareh, Jusoh, & Sheikh-Osman, 2008) addresses ambiguity issues in extracting concept and entity in NL test. They presented new technique that adapts possibility theory, fuzzy set and knowledge about the context to lexical semantics to resolve the ambiguity problem. The research partitioned into two parts of analysis (syntactic analysis and semantic analysis).

(Fabbrini, M.Fusani, S.Gnesi, & G.Lami, 2001; Gnesi, Lami, Trentanni, Fabbrini, & Fusani, 2005) developed an automated tool in detecting linguistic defects in NL requirements by following rules stated in the constructed quality model. In handling ambiguity that occurs in RS, The research stated four main attributes of ambiguity (vagueness, subjectivity, optionality and weakness) and three main attributes of understandability (multiplicity, implicity and unexplanation).

Apart from previous researches that attempt to minimize ambiguity, there is still limited researches focus on ambiguity that occurs in Malay sentences. Although requirements are commonly written in English language, there are still companies that writes system and user requirements in Malay language such as some departments government bodies and local-based SME software companies. We believe there is a need to analyze and develop a suitable technique that is best to reduce ambiguity occurs in Malay textual requirements.

Malay ambiguity

Malay language is used widely in documents throughout Malaysia, Indonesia and Brunei. Many small software companies in Malaysia still use Malay language in their requirements. Malay language structure is very much different from other language such as English that is widely used in the world. For example, the position of verb agreements, the use of articles and comparative discourse are different. Meanings in Malay sentence may vary even though they have the same words, phrases or even sentences. The different meaning in Malay sentences may include 1) Different sentence with same meaning, 2) Different verbs with same object and subject, 3) Similar verbs which do not have same meaning, 4) Different verbs with same meaning, 5) Different style and mechanics (Aziz, Ahmad, Ghani, & Mahmod2, 2006).

Most of the Malay language structure is dissimilar with English language or other language such as Arabic, Chinese, and Japanese etc. Example, pronouns in Malay is different from English; antecedent '*he/she'* clearly define the gender of a person, while in Malay, *'dia'* , '*baginda*' could refer to male or female (Noor, Noah, Aziz, & Hamzah, 2010a). 'Part of speech' sentence tagging is important because it is one of the common procedures for morphological analysis. The sentences and phrases needs to be parsed into its root form in order to detect its intended meaning (Ahmad-Nazri, Shamsuddin, & Abu-Bakar, 2008; Al-Fawareh et al., 2008). However, unlike western languages, some words in Malay can be tagged into more than one grammatical class. There are words in Malay that seem to correspond to "verbs" and they are also "adverb", "nouns" can be "prepositions". For example, the word "telefon" can be categorized under noun and also verb class. The word "boleh" can also be categorized under "noun" and also "verb" (Knowles & Mohd.Don, 2003). Previous research (Rojas & Sliesarieva, 2010) suggested the potential ambiguous word groups can be vague adverbs usually modifying nouns(such as acceptable, high, low, fast, etc), non deterministic adverbs usually modifying verbs (such as continually, periodically, regularly etc, general verbs that reflects inaccurate description (such as process, monitor, support etc), non deterministic constructs such as and/or, any, not limited to etc).

A sentence must be checked against its' grammatical structure considering morphological and syntactical aspect of a sentence. (Hirst, 1987) stated that

three grammatical elements that touches morphological processes in Malay language are affixation, reduplication and compounding. Reduplication of nouns triggers semantic category of heterogeneity or indefinite plural while reduplication of verbal results in one semantic feature such as repetition, continuity, habituality intensity, extensiveness and resemblance. Many researches seem to agree that morphological analysis in one of the important element to disambiguate phrases. He categorized 5 classes of Malay grammar elements which are nominals, verbals, auxiliaries, adverbials and particles. The word class however depends on its affix and affix deletion to get the root form of the word.

RESEARCH APPROACH

The NLP components that we make use of are parsing method, morphological analysis, lexical analysis and syntactic analysis. We have chosen to adopt a rule-based POS tagger method (Alfred, Mujat, & Obit, 2013) for our chunking and POS tagging activities. It uses 18 main tag set although there are many versions of Malay tag set. We decided to adopt this method as this is more suitable and relevant to our scope of samples. This method caters 18 tag set for Malay words categories. It also manages situation where a word has more than one word senses by providing word relation rules. Apart from that, it has guiding rules to assign POS tag to words in the event where affixes concern.

In order to extract the vague elements in a sentence, we use the hybrid method by combining two constructs; vague criterion and ambiguity attributes. We digest and gather relevant vagueness criteria from literatures that spread across discipline; language, biomedical and computational areas to come out with a model of vagueness. To ensure this criterion complies with the definition of requirement ambiguity, we then classify and categorize ambiguity into six attributes. To support the accuracy of the extracted vague concepts, we impose human expert verification to the lexicon to ensure we extract the right ones. As for the detecting method, we adopt string-matching and pattern-matching approach. It is where we compare the input sentences with our defined lexicon. In order to validate the method we proposed is significant in assisting users, we will compare the experimental result from automated prototype tool with manual human judgement method. A manual textual requirement of a system will be handed to relevant users to identify the ambiguity. Figure-1 articulates the conceptual diagram of our Malay Ambiguity Detection (MAD) approach.

PROPOSED CONCEPTUAL MODEL

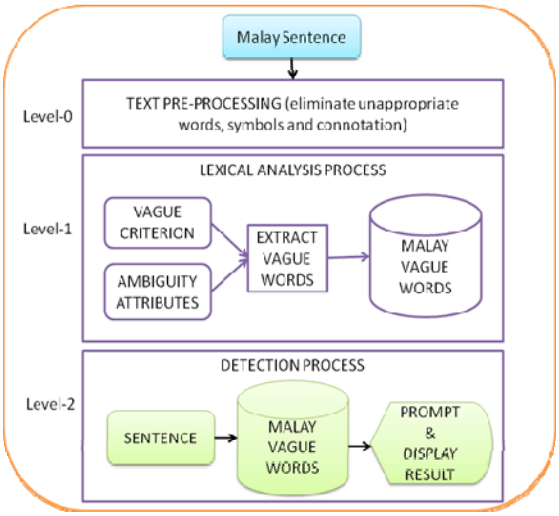

Figure-1. Conceptual model of MAD approach.

Figure-1 above depicted the conceptual model for our MAD solution. It consists of three levels; Level 0 for the text processing part, Level 1 for the Identification of potential vague Malay words processes and Level 3 for the detection process. What makes our model different from other researches is we include the vagueness element in our model due to the fact that vagueness has a significant effect on ambiguous sentences.

Level-0: Text pre-processing

The sentences from a sample of 10 Malay RS are pre-processed by eliminating the inappropriate words, symbols and connotations. The examples are double words such as 'rekod-rekod', 'bersama-sama', 'fungsi-fungsi' and etc. Words that come with prefix and suffix will go through a morphological process leaving only root words. Symbols are such as ", `, ?, ! and etc. The cleansed words are then saved into an Excel format repository for further processing.

Level-1: Lexical analysis process

Lexical Analysis Process is a level that identifies potential Malay vague words commonly used in samples of Malay RS. The identified words are kept in a MSSQL database. There are three elements involved at this level which are A. Vague Words Criterion, B. Ambiguity Attributes and C. Malay Ambiguous Words (MAW) lexicon.

A. Vague criterion

A dictionary of 100 ambiguous Arab words that has been developed, takes into consideration more than 10 word senses as the criteria (Merhbene, Zouaghi, & Zrigui, 2010). These senses were extracted from the Arab dictionary. Chantree et al. extracted ambiguous sentences indicate coordination ambiguity and developed ambiguity threshold to set the ambiguity benchmark (Chantree,

Roeck, Nuseibeh, & Willis, 2006). Amongst the factors involved in making sure readers understand what a sentence means are sentence length, ambiguous adjectives, adverbs and passive verbs (Ormandjieva, Hussain, & Kosseim, 2007). A list of high potential English ambiguous words has been constructed in an Ambiguity Technical Report as a guideline to avoid ambiguous sentence (Bender, 2003). Tjong et al. developed rules for clearer sentences in an attempt to avoid ambiguities (S. F. Tjong et al., 2007). These research proof that to begin an investigation to disambiguate an ambiguous sentence, one has to start by determining and identifying the vague words. These vague words could bring misconception and misinterpretation to the readers. As for the writers, they usually are not aware that they are even writing an ambiguous sentence in the first place. Through previous literatures as guidelines, we have tabled out a criterion of potentially ambiguous words that acts as guidelines to extract the poor words as in Table-1.

B. **Requirement ambiguity attribute** (RAA)

We constructed a set of Ambiguity Attributes adopted from previous quality attributes' literatures. We customized and modified taken into account only those of which most suitable for Malay words. It consists of six attributes as Table-1 below. The ambiguous Malay words are extracted based on these attributes from working RSs and some have been translated from English using Dwibahasa Kamus Oxford Fajar (Hawkins, 2007). Some of the word class attribute's words were extracted from Kamus Komprehensif Bahasa Melayu (Othman, 2005) for their POS.

Level-2: Detection process

Ambiguity detection activities occurs at this stage. Sentences from Malay RS are chunked, tokenized and tagged with their appropriate POS. The parsed words are being compared with the collection of identified Malay vague words. When any words matched the lexicon, system will prompt that the sentence is potentially ambiguous due to pre-identified reason. Although system has detected ambiguity from the sentencex, it is still human's decision whether to accept the prompted potential fault or to reject.

Table-1. Mapping of RAA and vague criterion description.

RAA	Vague criterion	DESC.	Example of vague Malay words
Implicit (Bender, 2003; Gnesi et al., 2005)	General	A form of vagueness that refers to subject or object in the sentence is generic rather than specific.	Efisien, mudah, pantas
	Subjective	A form of vagueness that refers to personal opinion or feeling	Mungkin, berkemungkinan
	Unquantifiable	A form of vagueness that refers to word that does not reflect any quantifiable measure	Termasuk,, patut
Connectives (Chantree et al., 2005, 2006; Gnesi et al., 2005)	Adjective	A form of vagueness that refers to words belonging to one of the major form classes in any of numerous languages and typically serving as a modifier of a noun to denote a quality of the thing named, to indicate its quantity or extent, or to specify a thing as distinct from something else	Segera, lengkap,seperti
	Adverb	A form of vagueness that refers to words belonging to one of the major form classes in any of the numerous languages, typically serving as a modifier of a verb, an adjective, another adverb, a preposition, a phrase, a clause, or a sentence, expressing some relation of manner or quality, place, time, degree, number, cause, opposition, affirmation, or denial, and in English also serving to connect and to express comment on clause content	Besar, kecil
	Verb	A form of vagueness that refers to wordw that characteristically is the grammatical centre of a predicate and expresses an act, occurrence, or mode of being, that in various languages is inflected for agreement with the subject, for tense, for voice, for mood, or for aspect, and that typically has rather full descriptive meaning and characterizing quality but is sometimes nearly devoid of these especially when	Ditetapkan, dikawal

		used as an auxiliary or linking verb	
	Function Word	A form of vagueness that refers to function words which typically combines with a noun phrase to form a phrase which usually expresses a modification or predication	Di, kepada, dalam, itu, ini
	Dangling Else	A form of vagueness that refers to the requirement has no other exit when one case is not met (Exception case)	Boleh, berkenaan
Temporal (Gnesi et al., 2005; Sc & Eng, 2002)	Indefinite timing or duration	A form of vagueness that refers to words that has time/duration type that invites multiple interpretation. Un-boundary timing or duration	Mingguan, bulanan, dari semasa ke semasa
Referential (Gnesi et al., 2005; Karimah et al., 2012; N, Abd, Azman, & Noah, 2011; Noor, Noah, Aziz, & Hamzah, 2010b; Sc & Eng, 2002)	Multiple objects/ Anaphora	A form of vagueness that refers to a subject that points to one or more objects in a sentence.	Berikut, seperti, salah satu, diantara
Domain-Specific Variable(Bender, 2003)		A form of vagueness that refers to domain specific variables invites vague interpretation and understanding. Too generic.	Aplikasi itu, pangkalan data, proses, data
Weakness (Fabbrini et al., 2001)		Sentence that contains weak main verb	dianggarkan

RESEARCH METHODOLOGY

Due to the limited establishment of ambiguity detection method for Malay language sentences, this research adopts other languages' methods and techniques with modification and customizations that best suit the research scope. The core of the research is the process of identification for potential vague Malay words which commonly used in the working Malay RS. The samples of RS were collected from few software development companies. We combined appropriate vague and ambiguous criterion adapted from previous literatures to construct a set of Malay vague criterion of our own (Haron & Abdul Ghani, 2014). We propose to proceed with the research according to the following steps:

- We conducted a literature review on the ambiguity detection technique from the perspective of computational linguistic area, compared and evaluated their approaches as well as methods. We also go across multi discipline in getting the techniques for other languages too to see the methods used.

- We identify vague criteria across discipline and constructed our model of vagueness to be the base of our vague concept extraction (Bennet, n.d.; Braun & Sider, 2007).
- We classify and categorize requirement ambiguity attributes to be mapped with the identified vague criteria (Bender, 2003; Fabbrini et al., 2001; Gnesi et al., 2005; Shiffman, 2005).

- We collected a set of Malay requirement samples from industry in the domain of medical and university information system to be as the training set.

- We extracted the vague concepts from the samples of requirements based on the model of vagueness and get it verified by the human experts. The Model of Vagueness and the descriptions for extracting vague concepts is shown in Figure-2 and Table-1 accordingly (Bender, 2003; Gnesi et al., 2005; Shiffman, 2005).

- We developed a POS tagger tool based on the adoption of the rule-based Malay POS tagger technique for the parse and tag activities (Alfred et al., 2013; Mohamad, Omar, Aziz, & Rahman, 2011).

- We developed an initial MAD prototype tool as the experimental tool based on our designed conceptual model.

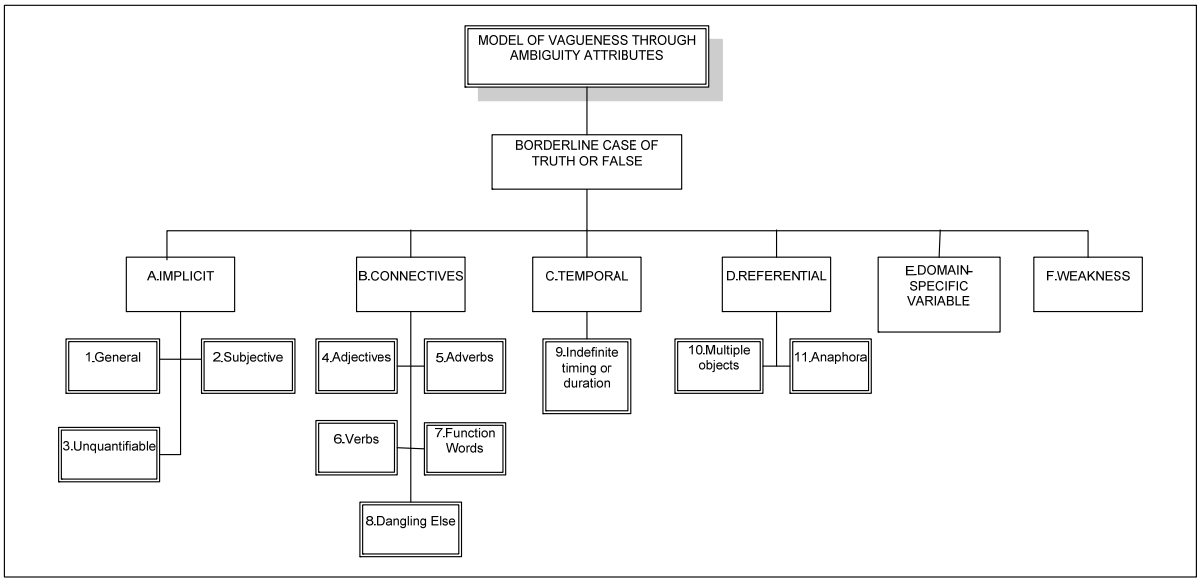

Figure-2. Model of vagueness through requirement ambiguity attributes (RAA) mapping.

RESEARCH RESULT AND PROGRESS

Based on the conceptual model in Figure-1, we developed an automated ambiguity detection prototype tool using Php and MSSQL. We constructed six Requirement Ambiguity Attributes (RAA) with the purpose of as the quality guidelines to extract the potential vague concepts. These two constructs are combined into a model of vagueness. A model of vagueness that consists of 11 vague criterions has been designed as shown in Figure-2. Based on this model of vagueness, a collection of 120 potential Malay vague words and 135 Malay sentences has been identified and extracted from the training set. It has undergone first level of verification, and based on the feedbacks, we did some modification accordingly. This collection of vague words is kept in a repository known as Malay Vague Words (MAW) corpus. Figure-3 shows a difference of MAW before and after the verification activities.

Table-2 below depicted the result after the second stage verification made by the expert. It shows the top three categories that contributes to high potential of vague words. The attributes are Ambiguous Adjective (ADJ) followed by Unquantifiable Boundary (UQB) and Referential (REF). Examples of vague Malay words that falls under the top three categories is shown in Table-3.

Our next step is the evaluation stage on the prototype tool. This is a way of validating our constructed models (Conceptual Model, RAA, Model of Vagueness). We will be using 10% of the collected industrial Malay requirements and the verified MAW lexicons in the evaluation part. We are in the progress of developing a System Requirement Specification in Malay language (M-SRS) that has the elements of potential ambiguous words as our evaluation instrument.

Table-2. Total and percentage MAW by ambiguity attributes.

CAT	TOT	%	CAT	TOT	%
IMP	19	14.1	WV	6	4.4
MS	4	3.0	PO	7	5.2
ADJ	41	30.4	REF	41	30.4
ADV	10	7.4	TEMP	10	7.4
UQB	35	25.9	SF	0	0
DE	25	18.5	DST	4	3.0

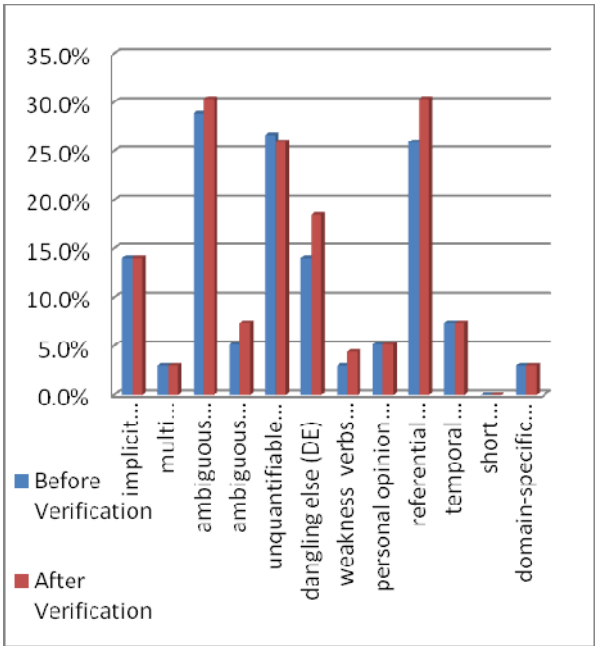

Figure-3. Comparison of MAW before and after expert's verification.

Table-3. Examples of identified maw with sentences for the top three vague criterion.

Category	MAW	Example of sentences
ADJ	segera	Wakil Jabatan perlu menjawab soalan tersebut dangan kadar segera ataupun memberi kepada keesokan hari sekiranya memerlukan siatan atau kajian lanjut
	ketat	Capaian fungsi atau maklumat adalah ditetapkan dan dikawal ketat, di mana pengguna hanya boleh mencapai apa yang dibenarkan sahaja
REF	aplikasi	Ini bermakna, aplikasi luar daripada pelayan tidak boleh membuat query ke dalam pangkalan data sistem
	Dari masa ke semasa	pengenalpastian dari masa ke semasa amat diperlukan
UQB	segala	mengandungi segala urusan yang berkaitan
	kesesuaian	jawapan akan dikemaskini mengikut kesesuaian semasa persidangan kelak

The identified vague words should be avoided from being used by the writer in an attempt to minimize the possibility of making an ambiguous sentence. This research will proceed with evaluation phase where we will test our prototype with the identified vague words in order to auto-detect potential ambiguous sentence and message of faults will be prompted. Our future work will be the results of the Level-2 components from the proposed conceptual model.

CONCLUSIONS

Ambiguity is a widely known error occurs in NL based documents. 60% of software development errors originated from the misunderstanding in RS (Shukur, Zin, Ban, & Ping, 2006). More often than not, stakeholders and users does not aware the occurrences of ambiguity in the sentences they read or write (Chantree et al., 2005; Erik Kamsties & Paech, 2000; F. Tjong & Berry, 2008). One of the main sources of ambiguity is the usage of vague words. Although vagueness is not ambiguity but vagueness influence the occurrence of ambiguity. This study is making an attempt to identify and emphasize on the significant influences of vague words could make to a sentence. Malay words that have the feature of unquantifiable and in between of truth boundary should be avoided.

REFERENCES

Ahmad-Nazri, M. Z., Shamsuddin, S. M., & Abu-Bakar, A. (2008). An Exploratory Study on Malay Processing Tool for Acquisition of Taxonomy Using FCA. Eighth International Conference on Intelligent Systems Design and Applications. IEEE. doi:10.11.1109/ISDA.2008.254

Al-Fawareh, H. M., Jusoh, S., & Sheikh-Osman, W. R. (2008). Ambiguity in Text Mining. In Proceedings of the International Conference on Computer and Communication Engineering 2008 (pp. 1172–1176). KL, Malaysia: IEEE.

Alfred, R., Mujat, A., & Obit, J. H. (2013). A Ruled-Based Part of Speech (RPOS) Tagger for Malay Text Articles, 50–59.

Ambriola, V., & Gervasi, V. (1997). Processing Natural Language Requirements.

Aziz, M. J. A., Ahmad, F., Ghani, A. A. A., & Mahmod2, R. (2006). Pola Grammar Technique for Grammatical Relation Extraction in Malay Language. Malaysian Journal of Computer Science.

Bender, R. (2003). The Ambiguity Review Process. Bender RBT Inc.

Bennet, B. (n.d.). Modal Semantics for Knowledge Bases Dealing with Vague Concepts.

Berry, M. D., Kamsties, E., Krieger, M. M., & Lee, W. L. S. & T. (2003). From Contract Drafting to Software Specification: Linguistic Sources of Ambiguity.

Braun, D., & Sider, T. (2007). Vague, So Untrue. NOUS, 2, 133–157.

Burg, J. F. M. (1989). Linguistic Instruments in Requirement Engineering. IOS Press Inc.

Chantree, F. J., Nuseibeh, B., Roeck, A. De, & Willis, A. (2005). Nocuous Ambiguities in Requirement Specifications. UK: The Open University.

Chantree, F. J., Roeck, A. de, Nuseibeh, B., & Willis, A. (2006). Identifying Nocuous Ambiguity in Natural

Language Requirements. Faculty of Maths and Computing. The Open University, UK.

Fabbrini, F., M.Fusani, S.Gnesi, & G.Lami. (2001). The Linguistic Approach to the Natural Language Requirements Quality: Benefit of the use of an Automatic Tool. In Software Engineering Workshop, 2001. Proceedings. 26th Annual NASA Goddard (pp. 95–105). Greenbelt, MD , USA : IEEE Xplore.

Gnesi, S., Lami, G., Trentanni, G., Fabbrini, F., & Fusani, M. (2005). An Automatic Tool for the Analysis of Natural Language Requirements. International Journal of Computer Systems Science and Engineering.

Grenat, M. H., & Taher, M. M. (2008). On a Translation of Structural Ambiguity. Al-Satil Journal, 9–19.

Haron, H., & Abdul Ghani, A. A. (2014). A Method to identify potential ambiguous Malay words through ambiguity attributes mapping: An exploratory study (pp. 1–8).

Hawkins, J. M. (2007). Kamus Dwibahasa Oxford Fajar. (J. M. Hawkins, Ed.)Kamus Dwibahasa Oxford Fajar. Oxford Fajar Sdn Bhd.

Hirst, G. (1987). Semantic iterpretation and the resolution of ambiguity. Cambridge University Press.

Huyck, C. R., & Abbas, F. (n.d.). Natural Language Processing and Requirements Engineering: A Linguistics Perspective.

Kamsties, E. (2005). Understanding Ambiguity in Requirements Engineering. In Engineering and Managing Software Requirements. Springer-Verlag Berlin Heidelberg.

Kamsties, E., & Paech, B. (2000). Taming Ambiguity in Natural Language Requirements. In International Conference on System and Software Engineering and their Applications (pp. 1–8). Paris, France.

Karimah, N., Noor, M., Noah, S. A., Juzaiddin, M., Aziz, A., & Hamzah, M. P. (2012). Malay Anaphor and Antecedent Candidate Identification : A Proposed Solution. Springer-Verlad Berlin Herdelberg 2012, 141–151.

Knowles, G., & Mohd.Don, Z. (2003). Tagging a corpus of Malay texts and coping with "syntactic drift," 422–488.

Merhbene, L., Zouaghi, A., & Zrigui, M. (2010). Ambiguous Arabic Words Disambiguation. 2010 11th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, 157–164. doi:10.1109/SNPD.2010.32

Mohamad, H., Omar, N., Aziz, M. J. A., & Rahman, S. A. (2011). Statistical Malay Dependency Parser for Knowledge Acquisition Based on Word Dependency Relation. In

Pacific Association for Computational Linguistics (PACLING 2011). Elsevier Ltd.

N, N. K. M., Abd, M. J., Azman, S., & Noah, M. (2011). " nya " as anaphoric word : A proposed solution, (June), 249–254.

Noor, N. K. M., Noah, S. A., Aziz, M. J. A., & Hamzah, M. P. (2010a). Anaphora Resolution of Malay Text: Issues and Proposed Solution Model. 2010 International Conference on Asian Language Processing. Harbin: IEEE.

Noor, N. K. M., Noah, S. A., Aziz, M. J. A., & Hamzah, M. P. (2010b). Anaphora Resolution of Malay Text: Issues and Proposed Solution Model. 2010 International Conference on Asian Language Processing, 174–177. doi:10.1109/IALP.2010.80

Nuseibeh, B., Easterbrook, S., & Russo, A. (2001). Making Inconsistency Respectable in Software Development.

Ormandjieva, O., Hussain, I., & Kosseim, L. (2007). Toward a text classification system for the quality assessment of software requirements written in natural language. Fourth International Workshop on Software Quality Assurance in Conjunction with the 6th ESEC/FSE Joint Meeting - SOQUA '07, 39. doi:10.1145/1295074.1295082

Othman, A. (2005). Kamus Komprehensif Bahasa Melayu. (N. S. Karim, Ed.)Kamus Komprehensif Bahasa Melayu. Oxford Fajar Sdn Bhd.

Rojas, A. B., & Sliesarieva, G. B. (2010). Automated detection of language issues affecting accuracy, ambiguity and verifiability in software requirements written in natural language. Proceedings of the NAACL HLT 2010 Young Investigators Workshop on Computational Approaches to Languages of the Americas. Los Angeles, California: Association for Computational Linguistics.

Sc, J. K. D., & Eng, C. (2002). A Prototype Tool for Improving the Wording of Requirements. 12th Annual International Symposium of the NCOSE 2002.

Shiffman, R. N. (2005). A Model of Ambiguity and Vagueness in Clinical Practice Guideline Recommendations. In AMIA 2005 Symposium Proceedings

Shukur, Z., Zin, A. M., Ban, A., & Ping, see C. (2006). A Tool for Translating A Software Specification Written in Malay into A Formal in Z. Asian Journal of Information Technology, 5(3), 8.

Tjong, F., & Berry, D. M. (2008). Can Rules of Inference Resolve Coordination Ambiguity in Natural Language Requirements Specification? 11th Workshop on Requirements Engineering.