

## 1. INTRODUCTION

Al-Qur'an is a book of guidance which is the curriculum of life for humans in their lives to achieve happiness in the afterlife (Salim, 2015). The main language of the Quran is Arabic Language. A word can be derived into many words in Arabic Language. If we take a look in Arabic dictionary, the derivative words are mostly unlisted or uninserted in the basic words, that's why Arabic Language being difficult for beginner, because they have to know the basic words first, then find the meaning of the word intended in the part (index) of basic form words. Nahwu is a branch of Arabic Language about to arranging sentences in accordance with Arabic rules, both related to the location of words in a sentence or word condition (harakat and final form) in a sentence. Nahwu focuses on how to arrange words into a perfect sentence, both in terms of words arrangement or final change of each word in a sentence that known as i'rab.

Arabian concerns about standardization by establishing grammatical, as the symbol of culture and literature. Sometimes the standards of intellectuality by using common idioms regarded as personal image of individuals. Then, the development of Arabic Language science has adopted special knowledge about grammar and its derivatives. However, for speakers who the mother tongue isn't Arabic will find difficulties, such as the practice of Arabic Language in boarding schools. Study of the desire to master Arabic was raised, there are at least three things regarding to this learning problem. First, the widespread practice of i'rab using, it leads students to find difficulties. Mastery of Arabic isn't only about the ability to translate word into sentence by using i'rab. If we take a more explore, i'rab doesn't help us to learn using Arabic, but it's about the Arabic itself (Yoyo & Mukhlis, 2019).

In the sentence "ضَرَبَ بَكْرٌ زَيْدًا" which means "Bakr has hit Zaid", there is a syntactic relationship both word 'Bakr' as the subject who 'hitting', and 'Zaid' as the object that is 'beaten'. If the syntactic relationship is wrong, then the meaning of the sentence will be different. The perpetrator will be the victim, while the victim will become the perpetrator. For that reason, a parsing dependency for Arabic Quran was developed, in order to show syntactic relationships between words in a sentence.

Parsing is a descriptive linguistics exercise that involves breaking down the text into elements of speech with evidence of the form, function, and syntax of the relationship of each part so that the text can be well understood (ThoughtCo., 2019). Easy-first parsing is a transition based parsing which is a type of syntactic parsing algorithm (Nguyen et al., 2018). The Easy-First parsing algorithm performs work based on the difficulty level of the task. Thus, the easiest jobs (simplest) are executed first before finally executing the more difficult ones. Of course the most difficult work will be put at the very last step. This algorithm works in an easy first order, unlike other deterministic parsers which are limited by parsing order (left-to-right) (Goldberg & Elhadad, 2010).

Dependency Parsing aims to identify sentences and determine their syntactic structure (Medium, 2019). Dependency parsing is a necessary process for analyzing sentence structure, especially for the Arabic language of the Quran which has a rich language. In NLP, dependency parsing is very useful for some work cases involving text data such as question answering and information extraction. Dependency parsing is important to understand the natural language, that the performance may have direct effect on NLP applications (Pei et al., 2015). For example on question "Who is the current Indonesian Minister of Education?", dependency parsing will very helpful to determine the semantic relationship between words in the information source sentence by identifying their head-dependent, so that the answer will be found.

In the field of linguistics, dependency parsing is also an aspect that plays an important role, especially for languages that are rich in morphological aspects. Many researchers have experimented with dependency parsing in various languages. In Indonesian language there is a paper studied by Kamayani and Purwarianti. The paper describes Indonesian dependency

grammar based on Stanford Dependency Label. The parsing method used is deterministic parsing which is implemented in Prolog, by using Covington algorithm for free word order, it's separated the lexicon, grammar, and algorithm. The features of structures in words and dependency rules are represented using GULP (Graph Unification Logic Programming). The grammar is separated from parsing algorithm by the unification-based feature structures. The feature structures are notated and unification included in the grammar rule. The parsing won't be affected by the rule order (Kamayani & Purwarianti, 2011). This parser works for simple Indonesian sentences only, not yet able to handle sentences with complex clauses. Based on Stanford Dependencies, there are some rules can't be used in Indonesian. There is also a paper studied by (Marton et al., 2010). They delve the participation of various lexical features and inflectional dependency parsing to Arabic, using Columbia Arabic Treebank (CaTiB) based on the gold standard and prediction of POS conditions. The method used in this paper is also deterministic parsing. The paper said that the quality of parsing produced in gold is better than the quality of parsing in POS prediction conditions (Marton et al., 2010).

In this study, the method used is alike with (Kamayani & Purwarianti, 2011) and (Marton et al., 2010) work, which is using deterministic parsing method. It is said that the deterministic dependency parsing is one of the most reliable methods for parsing dependencies, providing that the deterministic parser is driven by qualified classifiers (Nivre, 2012). The gap in relevant studies (Marton et al., 2010) is that it was odd when a feature was used. Two correct dependency relations increase even though they do not use a feature, while one that uses them does not produce correct dependency relations. The parser used is MaltParser v1.3, a transition-based parser with input buffers and stacks and also determining the next state prediction with the SVM classifier. In this study, the deterministic parsing used is shift-reduce parsing, where the Easy-First parsing algorithm by Goldberg and Elhadad is implemented. The reason for choosing to use Easy-first parsing due to the sequence of processing is pretended by making easier decisions before the more difficult ones. While, at the decision point, shift-reduce parsers usually rely on syntactic information on the left hand side, because it is limited by the processing order from left to right. For all of this reason, it is hoped that it can solve the feature's involvement problem. In addition, this parsing algorithm greatly exceeds the left-to-right deterministic algorithm (Goldberg & Elhadad, 2010). Then, the evaluation was carried out using the labeled attachment score measurement method.

The purpose of this study is to make Arabic learners easier to understand and speak Arabic fluently. Of course, this language fluency is also followed by the correct use of Arabic grammar. Therefore, it is hoped that the dependency parsing for Arabic will make people who want to learn Arabic will easily understand sentence structure.