An Analysis of Prepositional-Phrase Attachment Disambiguation

Mohammed H. Hamdan, Imtiaz H. Khan
Department of Computer Science
King Abdul-Aziz University
Jeddah, 21577, Saudi Arabia
Email:Mhamdan0005@stu.kau.edu.sa, ihkhan@kau.edu.sa

ABSTRACT: *Prepositional-phrase (PP) attachment ambiguity is a pervasive problem in natural language processing, and at times it poses significant challenges to a computer system to resolve this ambiguity. In literature, different approaches have been proposed to address PP-attachment ambiguity, but to the best of our knowledge, there is no published work which surveys such approaches. This survey paper compares the standard methods that attempt to resolve PP-attachment ambiguities in natural language processing. We also provide a taxonomy of various ambiguities, which may arise at different levels during the language-processing task. There are two methods employed in natural language processing concerning new approaches: the first technique called the rule-based method and the second called statistical approach.*

1. Introduction

Natural Language Processing (NLP) aims at building computer systems to process human languages (like Arabic or English) automatically. Unlike programming languages, natural language poses significant challenges because the latter is notoriously ambiguous. Ambiguity is a characteristic of natural language, whereby the text can be interpreted in more than one different ways [1].

There are different types of ambiguities, like for example lexical ambiguity, structural ambiguity, semantic ambiguity and pragmatic ambiguity. These uncertainties arise at the various levels during the language processing, for example, lexical ambiguity arises at the word level, and structural ambiguity arises at the sentence level. Among these different types of ambiguities, the most common and widely studied ambiguity is structural ambiguity (also known as syntactic ambiguity), whereby a sequence of words can be grammatically structured in more than one way hence resulting in more than one different interpretation.

Psycholinguistic evidence suggests that in many cases such ambiguities confuse [2]. However, sometimes, even though

theoretically ambiguity can be presented in the phrase, most people may interpret the phrase in the same way. Therefore, the problem is how a computer system can determine the likelihood of different interpretations of a given expression and then be decided which analysis is most likely.

We are exploring the PP-attachment ambiguity problem whereby a prepositional may be attached to a preceding noun or verb. For example, in the expression 'Ali saw a man with the telescope,' one could be inclined to interpret it like a man having the telescope (i.e., PP-attached to the noun) or Ali used the telescope as a device to see the man (i.e., PP-attached to the verb).

To the best of our knowledge, no published literature attempts to explore the whole prevalent approaches to deal with prepositional phrase attachment ambiguities in natural language processing. Hence, in this paper, we endeavor to survey and analyze the common strategies that crack to resolve PP- attachment ambiguities in native language processing. We absorbed on presenting a state-of-the-art survey of the currently available tools for ambiguity resolution. The objective of this paper is to broadcast, divide and analyze the research work published in PP- attachment disambiguation.

There are different types of ambiguities which make NLP so hard, like lexical ambiguity, structural ambiguity, semantic ambiguity and pragmatic ambiguity.

Moreover, these ambiguities arise at the various levels during the language processing, for instance, lexical ambiguity occurs at the word level, whereas structural ambiguity arises at the sentence level. Among those multiple types of ambiguities, the most common and widely studied ambiguity is structural ambiguity (also known as syntactic ambiguity), whereby a sequence of words can be grammatically structured in more than one way hence resulting in more than one different interpretation.

NLP can be divided into two main categories, namely, Natural Language Understanding (NLU) and Natural Language Generation (NLG). Moreover, NLU copes with natural language as input for machines whatever its format is (text, sound, images, video, etc.). Then, any NLP application like 'Google' can represent and analyze such text.

Whereas, NLG considers natural language as an output from machines in its different formats. Which previously stored in machines' storages whereby computers able to deal with these formats and produce it to natural language either text-based or speech-based. The best-given example for NLG applications is automatic report generation and machine translation.

In the following sections, we discuss an earlier work that has been done on the ambiguity of natural language processing and more precisely focused on prepositional phrase attachment ambiguity. This contains previous studies on the types of ambiguity and different algorithms for PP-attachment disambiguation. In section 2, we present a taxonomy of different ambiguity classes and analyze each of these types succinctly. In section 3, we present and analyze PP-attachment ambiguity approaches, as we start with earlier approaches, followed by corpus-based methods, then, statistical approaches, Psycholinguistic motivated approaches, and other approaches. In section 4, we briefly discuss our work that provides a compile of different approaches to PP-attachment ambiguity resolution. Finally, we close with a short conclusion of this paper which provides an in-depth survey of the standard approaches used to resolve PP-attachment ambiguities in natural language processing.

2. Ambiguity Classes

Ambiguity is a pervasive problem throughout natural human languages, e.g., English, French, and Arabic, in both text-based and speech-based. Consequently, when ambiguity arose, it would make NLP so hard for machine learning in computational linguistics.

Furthermore, applications of artificial intelligence face a great challenging regarding processing human languages because of their ambiguities. The following diagram as shown in Figure 0.1 illustrates a taxonomy of different types of ambiguities, which may arise at different levels during the language-processing task.

2.1 Lexical Ambiguity
Lexical ambiguity is a widespread challenge in computational linguistic ambiguity. This kind of uncertainty can be broken down into two subproblems which are homonym and polysemy ambiguity [3][4]:
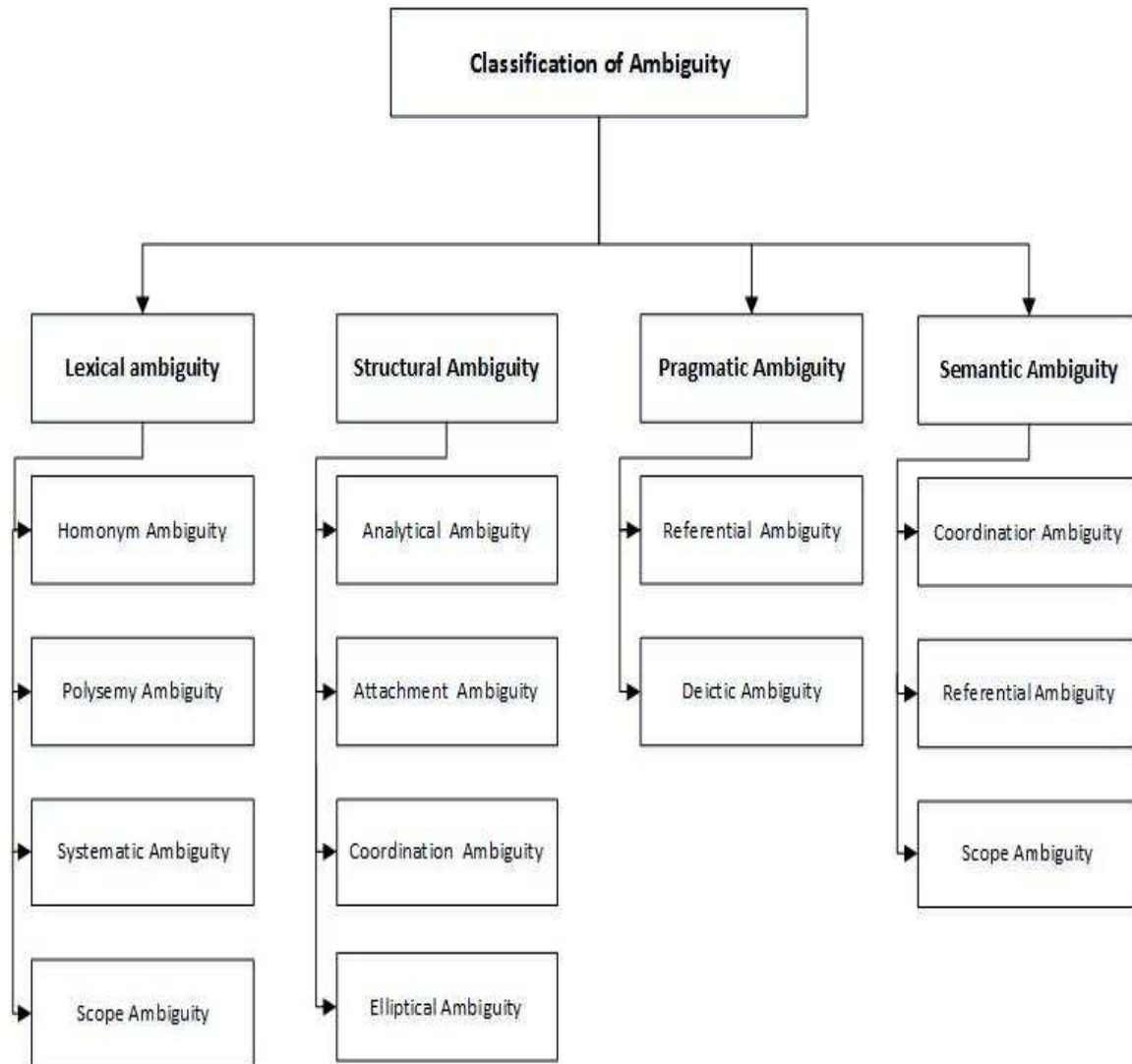
2.1.1 Homonym Ambiguity

Figure 1. A taxonomy of different ambiguity classes

Homonymy Ambiguity arises when two or more distinctive phrases or words share the same writing or sound. However, people can cope with numerous meanings and several etymologies efficiently compared with the machine, since it required a great effort. Whereas, scholars discover it to be too challenging to outline such consideration for this linguistic phenomenon. Therefore, it has led to some linguists to speak about a "polysemy paradox" [5]. An example of an ambiguous homonym that shared a similarity in both sound and spelling is "pen." "Pen" has two definitions: can mean both a holding area for animals and a writing instrument.

2.1.2 Polysemy Ambiguity
Polysemous Ambiguity occurs when a single vocabulary has quite a few associated senses, i.e., one etymology. For instance, the most regularly quoted [6] and discussed cases is "bank" which contains different interpretations while sharing a common etymology. Accordance to Longman's dictionary we found, the word "bank." Which means: [building, economic institution, personnel, any man or woman when enjoying a Monopoly, storage facility, row of comparable things, sloping and beside a physique of water, a large pile of snow, sand, earth, ridge and undersea elevation]. The meaning of these words is entirely based on the context.

2.1.3 Systematic Polysemy

Systematic polysemy is a kind of ambiguity continually arises at the polysemous vagueness which will be misclassified [7]. For example, (type - unit) uncertainty classes in this sentence "I like this pair of sunglasses" would possibly refer to either an individual sunglass or to a brand of sunglasses.

2.1.4 Scope Ambiguity
Scope ambiguity takes place when at least two or more equivalent expressions and quantifiers can overlap each other in distinctive approaches in the meaning of a sentence [8]. Consider this example "every woman loves a man." The more prominent meaning of this sentence is that for each woman, there is a man, and it is possible that every woman loves an extraordinary man. Nevertheless, the sentence also has a second feasible meaning, which says that there is one precise guy who is loved by using every woman. On the other hand, the sentence "I will give her a ring tomorrow" has two potential meaning. One interpretation may signify merely on an intention to call her, or it may be interpreted as a given promise of a gift of jewelry.

2.2 Syntactic Ambiguity
Syntactic ambiguity, which also called structural ambiguity, occurs when a sentence or phrase contains more than one parse. This category is, then, subdivided and classified into four types of uncertainty [9]. The subdivision of this ambiguity will be described in the next subsections.

2.2.1 Analytical Ambiguity
Analytical ambiguity is the reason behind the presence of an individual vocabulary that has two or more entirely distinct idea or thing needs to be represented [10]. Therefore, the structure of complex noun group may arise among various patterns of analytical ambiguity, e.g., the phrase "English Teacher" could have two different meanings. The first meaning maybe the teacher is from England, or the second one may be the teacher is from somewhere else, but he teaches English [11].

2.2.2 Attachment Ambiguity
Attachment ambiguity is the most public sorts of structural ambiguity. Moreover, the more frequently kinds of attachment ambiguity are PP-attachment and relative clause attachment.

A timeworn example for PP-attachment is the sentence "the policeman saw the thief with the telescope." It is difficult to find whether the PP (with the telescope) attached to the noun phrase (the thief) when the interpreted reader goes as the thief has a telescope who is seeing by the policeman.

On the other hand, it could be attached to the verb phrase (see) thereby the policeman had a telescope an instrument for seeing the thief [12]. Whereas relative clause attachment connected at least two phrases that composed a sentence.

2.2.3 Coordination Ambiguity
Coordination ambiguity considers the second common type of structural ambiguity in many languages, occurs in sentences which have conjunctional tools such as and, or, but. E.g., (black cats and dogs) [1].

The above example is ambiguity because the reader might have more than one interpretation. Thereby the black may refer only to the cats then it will read as (black cats) and dogs. On the other hand, the black can modify the dogs then it will read as cats and (black dogs).

Finally, the black can modify both cats and dogs. So, the adjective can be understood as black (cats and dogs). All these interoperations wholly based on the external modifier and readers.

2.2.4 Elliptical Ambiguity
An elliptical ambiguity arises because eliminating redundant words from sentences [13] in the Arabic language called Pro-Drop ambiguity.

However, it is not necessary to omit such word to cause such kind of elliptic ambiguity. E.g., the phrase "David knows a richer man than Bill." In this sentence there are two cases occurred: the first case, when there is no ellipsis, and the interpretation goes like this, there is a richer man than Bill who knows by David. On the other hand, when there is an ellipsis, and a verb "knows" was omitted then the interpretation goes like this, David knows a richer man than Bill knows. Consequently, a verb "knows" must come after "Bill."

Another example is "I bought and ate eggs and Danial vegetable." The reader may interpret this sentence based on three potential meanings: "bought and eat," "bought only" or "eat" only. Like the following: (1) I bought and ate eggs and Danial bought and ate a vegetable. (2) I bought and ate eggs and Danial bought a vegetable. (3) I bought and ate eggs, and Danial ate a vegetable.

It's clear that an ambiguous ellipsis can arise when there are more words need to be retrieved and put them in their correct place.

2.3 Pragmatic Ambiguity

Pragmatic ambiguity focused mainly on the relation between structural language and its context. Therefore, linguistic and philosophical sciences are on the whole roots of pragmatic research area [14].

In somehow, pragmatic ambiguity appears to be like a semantic ambiguity except the first targeted on the entirely context-based meanings while the latter centered on invariant meaning. These referential ambiguity and deictic ambiguity are the numerous sorts of pragmatics ambiguities.

2.3.1 Referential Ambiguity

A referential ambiguity occurs mostly in words, pronouns, and associated clauses which are triggered an uncertainty both using referring to their ascendant or ancestor phrases [15][16].

For instance, the sentence, "Justin met his neighbor who has a kind heart" contains a referential ambiguity. So, a clause phrase "who has a kind heart" can be referred either to the Justin or his neighbor entirely by the context.

To clarify the referential ambiguity, let's consider another example, in the sentence "If an infant does no longer thrive on raw milk, boil it" a referential ambiguity arose when the reader interpreted the "it" pronoun either to the "milk" or the "thrive."

2.3.2 Deictic Ambiguity

A deictic ambiguity occurs based on the result of the word or expression dependency on their possible meaning according to context [17]. The expression "I and you" arise the deictic ambiguity, because the pronoun "you" may express to the man or the woman.

Therefore, deictic expressions assist in setting up deictic roles which derive from the reality in day-to-day language behavior. So, deictic ambiguity occurs when the speaker addresses his utterance to different men or women and may additionally refer to himself or herself.

The sentence "every student thinks you are a genius" has a deictic ambiguity. As there are two potential interpretations either a quantifier "every" will include the pronoun "you" then deictic ambiguity occurs as an anaphoric expression to a student he/she. On the other hand, a quantifier "every" excludes the pronoun "You" from the sentence meaning.

2.4 Semantic Ambiguity

Semantic Ambiguity (SA) depends on the logical form when a single sentence contains neither lexical nor structural ambiguity [18]. Even though semantic ambiguity shares the similarity of structural, lexical, and pragmatic ambiguity, e.g., "Doctors hate it when children smoke" in this example the word "smoke" can be referred either to the children who are smoking or to someone while the children inhale their smoke.

However, SA has a different method to comprehend the context. Therefore, this type of ambiguity has been arisen by three different ambiguities, namely: coordination ambiguity, referential ambiguity, and scope ambiguity where those sorts of ambiguities are extracted from syntactic ambiguity, pragmatic ambiguity, and lexical ambiguity, respectively.

3. PP-Attachment Ambiguity Approaches

In this section, we present different methods that employed for resolving PP-attachment ambiguities in NLP perspective. Therefore, we explore an in-depth review work of PP-attachment disambiguation as presented in the following subsections.

3.1 Earlier Approaches

In the study [19], the authors tried to answer how to make parsers for computer programming languages. Therefore, the right

association approach has been undertaken. Consequently, they found out that the nearest item to the proposition should be attached as a noun phrase known as nominal attachment.

In this study [20], the authors expose in natural language processing which revealed many problems affected on NLP and shown prepositional phrase attachment ambiguity problems. Hence, it might be attacked some essential applications as Machine Translation and semantic role labeling which solved by maximum entropy approach and combined features.

Those features have already created by strong collocation properties in the different utilization of the target prepositional phrase attachment ambiguity. As it is a primary task of NLP, it can be solved using machine translation, and semantic role labeling. Therefore, they addressed the problem based on maximum entropy and combined features developed in the context of preposition sense disambiguation for semantic role labeling.

The elements of disambiguation system have been broken up into three categories: (1) Collocation features [21] are stimulated with the support of one-sense-per-collocation heuristic proposed through exploring strong collocation properties on different senses of the target preposition. (2) Syntactic aspects are proven the recursive syntactic of the target preposition. (3) Semantic position often relies upon on what features a prepositional phrase have; then they especially designed the points of each verbal and nominal propositions based on semantic information.

Unlike the previous study, in this study [22], the researchers proved the ability of reusable an entropy framework function. Which generalized to apply to all-natural language processing tasks.

Moreover, maximum entropy model was created as the opposite of likelihood method based on conditional probability. Furthermore, they built their dataset Treebank, Wall Street Journal (WSJ) from IBM company and University of Pennsylvania. 81.6% was obtained with an accuracy of resolution.

Authors used and investigated in the study [23] two different semantic role inventories. Meanwhile, those roles extracted their advantages from large annotated corpora. Whereas, the problem was a preposition classified on semantic roles for prepositions based on word-sense disambiguation and applied straightforward approach.

Furthermore, lexical associations used for relieving WSD features. The two performance accuracies 78.5%, 70.3% were gained based on two different datasets which are used by Penn tree bank and frame net, respectively.

Nevertheless, based on Weka's proposed by [24] the combined of both Penn tree bank and frame net led higher accuracy of 86.1% than the previous study. Which implemented using the decision tree learner by [25] of Quinlan's.

In this study [26] the authors implemented the classifier to solve PP-attachment heuristics from syntactic structures assumes so that the unclear data can solve by the precise context of the test phase. Furthermore, an accuracy achieved 82% for English text.

Unlike people, computers faced difficulties to solve such kind of ambiguity because the latter requires more knowledge and details about particular words. Let's consider the words "eat" and "drink" as known by the people. So, they have their related meanings based on their previous knowledge. However, computers need to feed them all those knowledge about words through huge dataset (corpora).

In contrast, when people don't have prior experience of words then they can't interpret the Scripture in context as same as computers when aren't fed by datasets. This work has done a classification task due to the problem in the above study. Furthermore, many studies, besides the primary goal of this study, are to predict the noun phrase or the verb phrase.

Correspondingly, heuristic extraction applied in this study based on chunked sentence either (VP, PP, NP) or (NP, PP, NP). So, the unsupervised algorithm used in this experiment which extracted the left side of a target preposition and then detected whether the PP attached either verbal or nominal.

Moreover, the datasets extracted from both the raw text of Linguistic Data Consortium of Spanish and annotated heuristic extracted sentences form Wall Street Journal though they found out an accuracy performance of 81.9% and 69%, respectively.

Authors in this study [27] focused on a word sense disambiguation problem using two powerful methods, namely, the semantic relations and textual glosses. Then, provided information by WordNet. While approaches of inferential heuristic on WordNet, they applied heuristic 831 times in Wordnet.

Furthermore, they found that noun phrase has more degree attachment to prepositions than verb phrase. Two types of heuristic rules were applied based on one verb or two verbs to detect PP-attachment. The dataset used in the study is the Wall Street Journal (WSJ) which revised to exclude the unambiguous classes. Consequently, the remaining dataset applied and the performance accuracy of 72.3% disambiguation rates.

In discourse level strategies a pragmatic ambiguity can be left unresolved because they don't affect the meaning of an utterance.

However, in this paper [28] the authors offered a unified account of the exact PP-attachment behavior. Then they concluded that the resolution of pragmatic of PP-attachment ambiguity is necessary for language understanding based on statistical analysis and a preliminary algorithm. Finally, the result parsing of a large text yielded 78% correct parses.

3.2 Corpus-based Approach
In this study [29], authors used a new method called Support Vector Machine for learning model to solve PP-attachment ambiguity based on sophisticated and semantic features of each preposition. Although, they faced the limit distance for two concepts to be matched and most of the words are semantically ambiguous unless disambiguated is hard to establish gaps between them.

Therefore, authors proposed supervised learning algorithm and unsupervised similarity-based iterative algorithm. Furthermore, they modified Quinlan's ID3 algorithm based on decision tree induction and combined the path between two nodes. As a result, their work study they found out accuracy of 88.1% based on the semantic dictionary for resolving the PP-attachment ambiguity.

Authors in this study [30] produced a new Pattern Dictionary of English Prepositions called (PDEP). So, preposition Patterns presented each pattern as a case of the template (Preposition) and complement sentence. British National Corpus was used and made a pre-process of the SemEval dataset, which contains 28,052 sentences. Moreover, tagging, parsing, and creating feature files for these sentences were taken less than 10 minutes with an equal time to upload the feature files.

In this study [31], the authors used an ontology-based dataset to analyze prepositional phrases. Also, authors mainly created F-logic queries containing about 883 questions; then, they found out with Broker system.

Dataset was converted to an ontology language. Then, they transformed the whole dataset into the ontology languages which evaluated on the system OntoBroker proposed by [32] queries in generic logical form. Accordingly, the accuracy achieved of 99.27% for a correct decision on PP attachment.

In this study [33], authors examined the meanings of the prepositional phrases. So, they exploited the use of over-identifies the semantic contexts and proposed two selection parsers for decoding the distinctive purposes. The location method relies on phrase meanings of prepositions as an alternative of contextual meaning.

Furthermore, machine translation (MT) equipment must be conserved primarily based on source language sentences or phrases. Although, there was a diverse variety of structure between the local language and foreign language.

Moreover, prepositions have been required semantic treatments as well for the disambiguation of the senses. In this study, the authors proposed two decision trees: first, is semantic analysis and, the second is the head method which determined that meaning of fist method.

Moreover, semantic analysis experimented to determine the implications of over from the grammatical features of the complements, and the other from the semantic character of the Heads. The first search for the sense of past should start with the decision tree by its addition.

Because most prepositional phrases utilized characteristics of both nouns and verbs as the supplements, denoted time over the

happen modifiers either verb phrase or noun phrase. The dataset of 500 sentences produced from British National Corpus, finally, they found a performance accuracy of 93.5% as a resulting of a system precision.

In this study [34], the authors presented an algorithm depends on tokenized corpora and used web-based collocations. The fact that, Arabic parsers and its PP-attachment approaches have lack of sufficient syntactic and lexical Arabic NLP resources.

Therefore, many problems that researchers faced challenging when applied corpus preprocessing of Arabic NLP resources and tools. Consequently, tokenization used preprocessing step for the current algorithm which used collocational association between the PP and it's candidate binders to solve PP-attachment ambiguity. This method reported performance rate near to 82% which is 10% better than the baseline model on Arabic Penn Treebank dataset.

In this study [35], authors focused mainly on the quadruple [VP, NP1, PP, NP2] sort of PP-attachment ambiguity. Moreover, supervised learning was applied on corpus-based of syntactically ambiguous and the algorithm was built a lattice of hierarchies. Furthermore, the hierarchy's elements of a quadruple input are combined and, then, two performance accuracies achieved of 87.23% and 90.53% for accuracy system without "of" preposition and with "of" preposition, respectively.

In this study [36], the authors faced some obstacles in their experiment which affected the efforts when they used a corpus and physical sciences articles. The experiment work study was broken up into two levels with different methods: (1) Back Propagation Networks approach that applied at the bottom for learning semantic constraints. (2) The top level's relaxation network method that practical on syntactic integration. Consequently, the noun phrase attachment was gained on the arbitrary length of representation.

Authors in this study [37] proved the most significant problem on a weakness of Arabic Treebank annotation guidelines due to the restricted annotators of Arabic Tree Bank 3 corpus. The dataset extracted from 'Annahar' newspaper based on inter-annotator agreement scores and paging scores between data producers and end users.

As a result, they found out near to 9% better enhancement for f-measure from 86.98% to 94.3% as an optimum value. To conclude, an accuracy of parsing Arabic treebank was lower than other languages such as English and Chinese Treebank, though they have the same size of data.

Authors in this study [38] undertook the shortage of Arabic text features: namely, POS tagging, base phrase (BP) Chunking, and tokenization. They used a supervised machine learning method which adopted from an English version.

Although Arabic text has its characteristics, the SML method which adopted from English work on Arabic text tasks. As a result, the found out presented of an Arabic Penn Treebank as a modern standard for Arabic corpus consists of 734 articles on various topics. The accurate performance of this study based on tokenization, POS tagging, and chunking task were 99%, 95%, and 92%, respectively.

In this study [39], authors addressed coordination ambiguity problem also called conjunction ambiguity. Which affected by an external modifier on the coordination word such as (and, or, but, etc.). Moreover, authors expressed the criteria for measuring a degree of ambiguity based on multiple judges to read the same ambiguous sentences, while this approach is suitable although it's expensive.

Moreover, authors utilized three different heuristic algorithms considering word appropriations in the general corpus: 1) Coordination matches which predicted coordination first among conjuncts coordinated. 2) Distributional similarity is an anticipated reading of coordination first when strong distributional similarity happened into two conjuncts. 3) Collocation frequency which is the last reading of predicted coordination that is considered when the collected modifier with the first conjunction more than with the second combination.

Furthermore, all above heuristics applied to British National Corpus dataset which is a useful technique for generating statistical information by sketch engine tool. All type syntactical relationship with distribution and similarity based on words and word sketches, eventually it led to a thesaurus.

Furthermore, they achieved three different results: 1) the coordination-matches heuristic accuracy of 43.6%, 64.3%, and 44.0%

were obtained on precision, recall, and f-measure, respectively. 2) The distributional similarity heuristic accuracy of 43.6%, 64.3%, and 44.0% achieved on precision, recall, and f-measure, respectively. 3) The collocation–frequency heuristic accuracy of 43.6%, 64.3%, and 44.0% achieved on precision, recall, and f-measure, respectively.

The authors of this study [40] developed the Sketch Engine tool which is useful corpus tool for solving NLP tasks like PP-attachment ambiguity. It provides over than 90 languages with 400 ready-to-use corpora as an input with appropriate linguistic markup.

Moreover, they generated amongst other things using word sketches based on Corpus Query Systems. So, around 8875 most frequent Czech words had been created based on sketches for the all those that achieved more than 1000 times in the corpus. Finally, the Sketch Engine has been built a thesaurus and Sketch differences.

In this study [41], authors extracted the data set from the web WWW which disambiguated and annotated by content creators considered as extensive training and its surface features.

Therefore, the ability and availability of large training dataset helped authors utilizing how languages work. Moreover, authors applied an N-gram statistical model which achieved a fitting optimal performance accuracy of 83.82% on a system which derived from variations reflected of both wordnet and the web search engine.

In this study [42] authors projected an approach of Word Dependency Distribution (WDD) with the utilize of both stationary distribution and Wordnet synset techniques. Which led to making PP-attachment predictions based on quadruple [VP, NP1, PP, NP2] input system. So, the method was applied to both Markov chain model and supervised method which are suitable techniques to decide the correct PP-attachment for a preposition phrase.

Therefore, authors provided two steps to solve this kind of problem: the first one was known as 'correct attachment' which mainly focused on statistical ambiguity based on supervised learning technique. On the other hand, 'selection attachment' which is relied on lattices of prepositional phrase attachment.

Furthermore, the PP-attachment dataset contained around 11000 ambiguous sentences. The dataset extracted from WSJ of the Penn Treebank. Eventually, this system was reported as performance accuracy of 87.54%.

Authors in the study [43] presented significant efforts to improve a combination of two methods: unsupervised learning and supervised learning. Thus, they attempted to improve the benefit from an extensive textual to develop this approach.

However, the problem appeared significantly whether there is something to be gained by combining unsupervised learning and supervised learning when scaling up both the seed corpus and the unlabeled corpus.

Therefore, an active supervised learning and unsupervised learning are the main two approaches which used on this studying with training data. Meanwhile, the dataset on this study contained 1-billion-words training corpus. That collected from many sources like WSJ, scientific articles of English.

This study [44] presented an unsupervised corpus-based approach to PP-attachment. Also, used a repeated process to be extracted from parsed corpus all training data. Furthermore, detecting attachment core for N2 among the two choices: V or N1 attachment in [V N1 PN2] quadruplet syntactic frame.

Moreover, collocation database used to control contextually similar eight words to the nouns and verbs in each quadruplet. Also, a large corpus used to create two datasets of the triples counts of the form (candidate, p, n2). As a result, a probabilistic score is assigned for verb attachment (V-score) and noun attachment (N-score).

Therefore, two cases of ambiguous sentences: the candidate that appeared within a short distance from the PP, but it is not the case in this study. On the other hand, the guaranteed incident that led to an unambiguous situation.

More details, two phases made for making a PP-attachment decision for a quadruplet [VP, NP, PP, NP] form. It was calculated an

average adverbial attachment score on VP, and second NP then replaced contextual similarity words, on the other side the calculation made based on an adjectival attachment score concerning. Then, returning the first NP with second NP on their similarity context consistently.

Finally, an attachment is determined by the combination of average scores for each attachment candidate and the candidate with the higher score that is beforehand selected. Consequently, authors found out a good result on the RRR dataset that improved construction algorithm and reflected the precision of the system with an accurate performance of 84.31%.

Authors in this study [45] designed board game paradigm which provided a natural speech data in controlled contexts for two experiments investigating prepositional phrase ambiguities.

The first approach is Prosodic which concerned with those elements of speech. Although, a Prosodic method is not individual phonetic segments it is the most critical units of speech. Consequently, authors achieved performance accuracy of 70% up to 78% on disambiguated sentences while they didn't prove that for low attachment ambiguous sentences based on context.

3.3 Statistical Approaches
Statistical parsers play an essential role in resolving prepositional phrase attachment ambiguity. Hence, we conducted many past studies that aim to tackle this PP-attachment ambiguity based on statistical methods.

In this study [46] authors predicted an incorrect attached to prepositional phrases as they considered the critical example of machine transaction applications. Therefore, a wrong PP-attachment occurs then led to severe errors in the translation.

Therefore, authors suggested overcoming this problem using classifier for path prediction which is set of four quadruple structured perception classifiers [VP, NP, PP, NP]. Moreover, the authors used Decomposition method-based inference algorithm to look at the problem of PP-attachment disambiguation.

The classical method for combinatorial optimization and applied to several inference problems in NLP. Meanwhile, an inference algorithm works on both parallel English and Hindi pair sentences. Furthermore, they used an iterative Coordinate Descent algorithm (called Algorithm 1) which calls the Project Algorithm (called Algorithm 2) until convergence.

Also, they used a dataset of Parallel Hindi corpora which created a corpus of 100 parallel sentences. As a result, they achieved an enhancement performance accuracy of 10% over the baseline system using Dual Decomposition system. Therefore, the accurate performance was increased from 54% to 64%. Finally, baseline system could apply to WH-clause attachment.

In study [47] authors created a system PPATTACH for disambiguation PP-attachment ambiguities using the utilization availabilities provided on WORDNET. Also, they considered ten presentations (of, with, in, as, by, on, to, from, at, for). Therefore, they used the selection algorithm with four quadruples [VP, NP, PP, NP] based on Verb-net and Word-net resource available on machine learning.

Moreover, the selection prepositional phrase ambiguity problem using Logistic Regression classification algorithm which implemented the decision procedures. Also, they used Weka2 software tool which provided by three essential features of each data: preprocessing, classification, and clustering.

They extracted around 23898 prepositional phrases from the Ratnaparkhi-1994 dataset. Finally, authors found out an accurate performance of 99% triggers to the noun phrase whereas they also found out for all prepositions an accuracy of 70.7% classification outperforming the PPATTACH system.

Authors in this study [48] provided an ambiguity solution which occurred in different places in NLP. Correspondingly, authors mainly focused on a performance of an anaphora resolution in machine translation (MT) systems which is poor.

Moreover, the anaphora resolution (AR) algorithm is a statistical and dynamic algorithm as well that makes using the facility of the least possible features. Therefore, the best advantage of this algorithm is required less human intervention from an NLP.

Though the main bottleneck is the absence of enough NLP resources, an algorithm enables utilizing collocational evidence,

Regency, bands, and searches for candidate antecedents in a 20-window size. Therefore, authors built golden standard set which contained in 5000 ambiguous that classified and allocated each category with a specific preposition to test their algorithm.

As a result, they found out the accurate performance to a golden standard with Anaphora Resolution algorithm achieved for precision, recall, and f-measure an accuracy of 78%, 100%, and 87% on Arabic Penn Tree Bank, respectively.

In this study [49] authors argued for different sources of information determining initial parsing decisions. That used with sentence fragment completions and self-paced reading experiments. As a result, they found out two decisions to solve syntactic ambiguity based on spending time for reading ambiguous preposition phrases.

Finally, they came up with two results: short time with fast reading VP-attached vague prepositional expressions. On the other hand, long time with low reading NP-attached. Also, the found out an account of 24% of the variance in reading times of the VP-attached PP.

In this study [50] authors expressed the differences between supervised learning and unsupervised learning methods. Then, applied the latter in their research in this study. Despite the lack of using the supervised process, it has powerful features that deal with corpora and can also be verified by humans with a precision of 84% on [VP, NP, PP] triple structural.

Therefore, the unsupervised method applied in this study to enhance performance accuracy of disambiguation prepositional phrase attachment ambiguity. Since they are utilizing an existing rule-based depends on the XIP parser proposed [51] dependency system for Fresh.

Furthermore, the authors utilized a statistical disambiguation algorithm and extracted dataset form the web-based lexical dependency for PP-attachment. Moreover, the authors showed the lack of regular MF2 on initial corpus dependencies which aimed to get a correct attachment of the parser for multiple headword attachments for the first analysis step.

Also, they extracted about 17.000 documents as the dataset from World Wide Web for every 869 queries 20 URLs retrieved using HTML tags. Similarly, the authors executed their experiment by using a set of Perl scripts which combined with UNIX command.

Furthermore, authors created and parsed all dependencies for PP-attachment on a new corpus contained 38.242.073 sentences and 1.368.903 for both words and phrases called web corpus. Because of the authors were mainly focused on quantity dependencies rather than quality, they produced a less reliable rule. Finally, at the end of the study, the good performance achieved with an accuracy of 83.21%.

In this study [52], authors proposed a solution for prepositional phrase attachment ambiguity based on revised a raw text. That annotated with part of speech tags and forms morphologically. Moreover, the authors used an unsupervised approach which relied on the heuristic algorithm.

Correspondingly, POS tagging and chunking the raw corpus based on preposition phrase and counts all [candidate, preposition, second noun phrase] triplets structure. Finally, as opposed to 'Stetina and Nagao.' Only unambiguous counts used in the raw corpus based on the hypothesis ambiguous attachment information. That solved ambiguous data test. Eventually, the authors reported on the RRR dataset an accurate performance of 83.7%.

The authors conducted this study [53], as a result of previous researches which applied the statistically lexical technique to solve prepositional phrase attachment ambiguity. Which considered the triplet [VP, NP, PP] structural PP-attachment ambiguity.

Moreover, the authors used Backed-off technique estimation which dealt with the issue of a prepositional phrase. Despite the rare of occurring multi PP-attachment, it could be attached to many sentences.

Furthermore, the authors provided the information about the most likely model which presented by the study [54]. That determined where is an attachment for prepositional phrases could be located using an estimated probability attachment.

Furthermore, they undertook two attachment events: the first one 1. and the second C assigns C assigned by two which C is indicated to the VP and NP2, respectively.

Meanwhile, they used Penn Treebank database which automatically extracted some quadruples. Furthermore, about 5% of the total data1014 out of 19963 ambiguous phrases among fourteen possible structures. It split into two datasets one is used for training while the other used for testing based on three prepositional phrases dataset. For PP1, extracted quadruples of about 5% of the total (1014/19963). Finally, they found out based on four lexical items an accurate performance of 84.5%.

3.4 Psycholinguistic Motivated Approaches

Authors in this study [55] verified a robust provided methods for resolving PP-attachment ambiguity problems. Moreover, authors presented two primary methods: the semantic method and the contextual information method. That relied on for determining pp-attachment ambiguity rather than lexical and syntactic information.

Therefore, authors of the study solved prepositional phrase attachment ambiguity based on radial basis function kernel with using of Support Vector Machine Learner (SVML) technique. Finally, they found out two similar results as an accurate performance of 92.85% and 93.62% according to FN and TB2 datasets, respectively.

In this study [56], authors applied the experimental methodology using the facilities of WordNet. That combined both an explored and some annotated properties of its senses and features. Moreover, they extended WordNet's features and studying the characterization of preposition behavior by including both Frame Net lexical units and Verb Net classes.

Also, the authors mainly focused on characteristics of a given text. Although they systematically extracted a quadruplet structural [VP, NP, PP, NP] dataset from the Penn Treebank Wall Street Journal corpus, they also manually generated constituency parses. Finally, achieved PP-attachment decisions for each ambiguous extracted preposition phrases.

Finally, authors used dataset consists of 27.937 quadruples vague PP-attachment sentences. Also, they divided the dataset into two sections: the first one is 20.801, whereas the second is 3.097 quadruples for both training and testing datasets, respectively. As a result, they found out based on test dataset an accurate performance precision of 88.4%.

Authors in these two studies [57] and [58] applied word sense disambiguation technique for solving prepositional phrase attachment ambiguity. Moreover, they proposed and implemented to the experimental review the most popular prepositions "for, from, with, and to." Which semantically collapsed of related senses prepositions.

Therefore, the semantically related prepositions sense was a combination of a cross preposition with its relations. Furthermore, each sense preposition expressed as a group of prepositions based on 'SemEval' corpus training. That used features of usability facilities label with its definitions as annotated and guidelines structures.

Consequently, those guidelines structures have helped the prepositions to validate the mapping of their senses relations. Finally, they found out an accurate performance of 76% and 74% based on both annotators result and original corpus, respectively.

Authors in this study [59] mainly focused on bilinguals at different levels of unbalanced comparison. Therefore, they tested the prediction of PP attachment ambiguity between a combination of two languages: first German foreign language. On the other hand, English native speakers concerning all levels.

Moreover, an ANOVA technique used to demonstrate the significant effect of attachment based on the entire dataset. Finally, the authors formalized the result based on all groups which categorized into four main classification categories: the first achieved an accuracy of 89.9% and 89.7% of VP attachment on both short intransitive and transitive VP attachment, respectively. On the other hand, the second achieved an efficiency of 8.6% and 5.1% of VP attachment on both quick intransitive and transitive DP attachment ambiguous, respectively.

Whereas, the third achieved an efficiency of 84.7% and 85.9% of VP attachment on both long intransitive and transitive VP attachment, respectively. Similarly, the forth attained an efficiency of 12.4% and 4.0% of VP attachment on both long intransitive and transitive DP attachment ambiguous, respectively.

In this study authors [60] projected a resolution of PP attachment ambiguity for Spanish to improve the performance of the general-purpose dependency grammar. Therefore, they applied the experimental study within the Free Ling open-source suite of

two powerful NLP's features: the first one is morphological analyzers, and the other is EsTxala tool.

Therefore, these devices also improved machine learning approach for the Spanish language. Finally, they found out the random baseline on the current 'EsTxala's tool' with an accurate performance of 66.3%.

Authors in this study [61] selected the best parser for PP-attachment disambiguation. Therefore, they used an Oracle-based rather than golden standard. Because of, an oracle has a robust feature to organize data, it has been replaced Bikel's parser.

In the experimental study, the authors converted the dataset RRR [62] into artificial sentences with quadruple [VP, NP, PP, NP] structured to be grammatically generated ambiguous sentences. Moreover, they manually provided an oracle dataset when there are no alternative attachments available on the ambiguous sentences.

Also, they also converted RRR dataset via documentation through each quadruple into a set of decisions based on Penn Treebank issued '0.5' which extracted. Finally, the size of the study dataset was 45,1563, and 3.097 sentences were extracted from PTB which used for training and testing, respectively.

Although, the dataset contained a combination of both ambiguous and unambiguous sentences the authors found out an accurate performance of 80.01%.

In this study [63], the authors examined a PP-attachment ambiguity which determined either as a verb phrase attached or noun phrase attached. Moreover, this study is the most frequent regarding spoken and written register. Then they investigated the possible effect of lexical factors on the frequency of VP vs. NP attachment. Which definiteness agreement on PP attachment as well as the influence of grammatical phenomenon to a unique language.

Additionally, they used the first filter manually to make sure that each sentence which included a triple [VP-NP-PP] structure and included the two online self-paced reading tasks. Finally, they found out an accurate performance of 61.4% on prepositional phrase attached to the former VP which counted more commonly than the NP.

As a result, authors proved that an accurate performance of native Greek speakers was much better than English, foreign language with achieved to the verb phrase attachment corresponding to the corpus data.

Authors in this study [64] mainly focused on assisting the performance of prepositional phrase attachment task with using of natural lexicalized dependency parsing strategies. Furthermore, they applied semi lexicalized approach and provided information on a lexical method which discriminatively trained from labeled data.

Whereas, the proposed plan in the experimental study also generate a lexical model learned by Expectation Maximization on unlabeled data. Whereby the latter method can perform better performance than labeled data.

As a result, they found out that when the study based on a discriminative model used labeled data, then the typical model structure works lower than learning from the unlabeled dataset.

Authors in this study [65] investigated the disambiguated for each significant closed of vocabulary. Moreover, they exploited the performance study which inspected on the whole dataset for each preposition.

Additionally, the authors used mostly the furthermost popular prepositions (of, in, from, with, to, for, on, at, into, and by) which performed using a sense heuristic algorithm that only considered the prepositions with high frequency. That carried a definite meaning used in everyday life. Also, they preprocessed and prepared the dataset in XML using both two methods: 'Senseval DTDs,' and 'Frame Net.'

Although they created the dataset from British National Corpus, which has been limited from other sources. Because of the ambiguous sentences for frame analysis that selected on Noun and Verb opened classes, the dataset should make available initiate prepositions.

Finally, the authors found out negatively correlation of -0.34 and 0.44 based on ordinary sense heuristic method. Also, they

observed results of each preposition as strongly negatively correlative based on two approaches: 'entropy and perplexity' using lexical sample disambiguation. As a result, the authors came up with a relative performance of [good, mediocre, and relatively weak] for each prepositions range.

In this study [66], authors exposed both semantic interpretation and syntactic analysis which required as an input for semantic analysis. Therefore, there were many parsing techniques raised to solve this issue.

Moreover, many algorithms used to parse sentences such as Top-down parsing which begun with the start symbol and seen what rules it figures in as the mother. Consequently, they looked at the daughters of these rules and decided whether the first one matches the next word in the input. If it does, do the same for the remaining daughters. If not, then go through the same loop again and this time matching the daughters of the rules already found with mothers of other regulations.

Finally, they applied Bottom-up parsing by matching words to the right-hand sides and of rules and, then, matching the resulting symbols to the right-hand sides of rules until they covered all dataset.

In this study [67], authors proved three steps which applied to improve performance and reduce the cost of machine translation. Therefore, they proposed the following procedures based on bi-directional English to Arabic languages: (1) analysis where the input sentence represented penalties which annotated by parser tree. (2) a transfer which allowed representation is moving from the Arabic language into another target language based on the tree to tree collection. (3) The generation which produced the target-language output in this phase. Finally, authors found out an accurate performance of 64% and 81% for both English-to-Arabic and Arabic-to-English, respectively.

In this study [68], authors solved PP-attachment ambiguity in both English and French sentences. Therefore, two kinds of rules were determined PP attachment the first is a low attachment which was considered as a noun phrase, and the second is a high attachment which recognized a verb phrase.

Moreover, the experimental study applied based on quadruple [VP, NP, PP, NP] structure with using the human sentence processing model (HSPM). Also, they used an ANOVA statistical tool for analyzing acquisition data from HSPM.

Although a structural level of the 9.441 English sentences and 1.437 sentences examined, they didn't find out an available ratio in the experimental study because of operational treatment of similarity.

In this study [69], authors applied a heuristic approach to help decision of selecting a specific preposition. Therefore, the preposition "of" allocated to the noun phrase in the whole dataset in the experimental study.

Moreover, a simple heuristic and plink function solved the problem by using the first nine PTB texts. Finally, they created a small sized corpus was contained in 69 sentences. Also, they maximized the brackets using Parseval technique. Whereas, they attempt to minimize any additional structure that 'Parseval' does not apply.

In this study [70], authors attempted to tackle the problem of a limitation approach on attachment relationships of quadruple [VP, NP, PP, NP] structure. Moreover, a new rule-based approach to PPA disambiguation is a transformation-based approach.

Additionally, the predication attachment based on many cases which merely initiated rules then automatically generated the contextual attachment using transformation approach. Also, the concrete problem in PP-attachment ambiguity solved in a quadruples form [VP, NP1, PP, NP2] structure.

As goes in this example "buy a car [PP with a steering wheel]." Thus, using transformation approach led automatically to predict proper attachment based on numerous possible contextual ciphers. Finally, they created around 12,766 ambiguous sentences with quadruples structured which extracted from a WSJ Penn Treebank. As a result, they found out an accurate performance of 76% and 81.8% without applied transformation rules and with enforced transformation rules, respectively.

Authors in this study [71] solved the problem of PP -attachment ambiguity based on a light approach. That built by the nearest neighbor algorithm. Moreover, they used the nearest neighbor algorithm with the assist of point-wise mutual information vector cosine and represented each word.

Therefore, the experimental study used a quadruple [VP, NP, PP, NP] structured for disambiguation PP-attachment ambiguity. Besides, then, the problem classified either be attached to the NP or VP based on the quadruple input ambiguous dataset.

Finally, they used the algorithm the nearest neighbors which worked via searching on the training dataset examples. And, then, whenever met the top-k nearest neighbors. So that, immediately determined and considered as attachment either VP or NP attachment based on a well-known classification of the nearest neighbors. As a result, they found out an accurate performance of 86.5% on the RRR dataset.

3.5 Other Approaches
In this study [72], the authors proposed a neural network based on recursively builds composite representations and word vector representations. Therefore, WVR applied in the experimental study. As a similar system, a linear classifier which accessed a wide range of features implemented and tested in the study.

Furthermore, semantic and syntactic are essential for solving preposition phrase attachment ambiguity. Moreover, two Arabic corpora considered for creating word vectors: 'arTenTen' and 'Arabic Gigaword arTenTen.'

Additionally, produced a large-scale corpus of automatically crawled web texts comprising 5.8 billion words after preprocessing the sub-corpus. Also, they extracted 130 million of the dataset which contained the word vectors training. Finally, an Arabic Gigaword corpus composed of newswire texts from several news agencies. As a result, accurate performance of 78.4% in this model as an accuracy.

In this study [73], the authors proposed a 'Winograd Schemas' approach WS. Which is one of the newest approaches that appeared recently for resolving preposition phrase attachment ambiguity? Moreover, WS approach solved a pair of sentences ambiguity which occurred either by one or two words create an ambiguity.

Additionally, World knowledge then reasoning for its solution are the position accomplishment for fixing its ambiguous sentences. Also, they used an AWS which a small reading comprehension test involving a single binary question. Finally, two examples illustrated (1) "the trophy would not fit in the brown suitcase because it was too big" followed by the question "what was too big?." The participant immediately goes throw answer with Answer" 0:" when the trophy Answer "1" the suitcase.

Then, a well-known example by Terry Winograd" (2) the city councilmen refused the demonstrators a permit because they [feared/advocated] violence". Therefore, if the word feared, then they seemingly refer to the city council; if it advocated, then they apparently apply to the demonstrators.

In this study author [56], mainly focused on solving and processing PP-attachment ambiguity problem which presented by quadruples [VP, NP1, PP, NP2] structured. Where VP, NP1, PP, and NP2 is a transitive verb, head noun of an object of VP, proposition, and the head noun of the subject matter of P, respectively.

Moreover, they applied nearest-neighbor classifier technique which achieved a better result than a study [29] with an accurate performance of 0.01% on different datasets. Also, they proved a vital to confine the filled condemnations in Penn Treebank.

Although They adopted the four-step classification procedure from the study [74] at each quadruple test, the training examples were previously sorted by a different vector composition technique. Consequently, the study relied on a set of the best-counted cases and when its company with equal divisions for NP and VP.

Therefore, the approach follows the next procedures when a specific class has most votes. Then, it will allocate to a quadruple. Finally, they developed a standard corpus from Penn Treebank which established by [75] and [22] for the binary classification problem which already described above.

Authors in this study [76] investigated and solved the problem of prepositional phrase (PP) attachment ambiguity. Moreover, in the survey, they considered an ambiguous example "The policeman watched the spy with binoculars."

With quadruple [VP, NP, PP, NP] structured. Therefore, the constructed ambiguous sentence as a watch, the spy, with, and

binoculars correspondingly to quadruple structured. Also, they implement an off-line binary technique to choose among Chinese whom their English a second language and a group of native English speaking for executing a sentence completion task. Finally, they carried out a referential context which examined both native and non-native speakers.

Although they divided their experimental study for each group into 30 participants for both Chinese learners and English native speakers using an off-line sentence completion task, they also, applied an online self-paced reading experiment to all participants as well.

Finally, they found out that Chinese learners based on on-line procession can be led stronger than native speaker regarding both semantic and pragmatic interpretation. As a result, they came up with non-ratio valuable which conducted to integrate bottom-up and top-down information during processing with achieved better proficiency for non-native speakers than native speakers.

In this study [77], the authors proposed alternative methods rather than the straightforward alphabetic list approach to reducing the time complexity for user searching. Moreover, the mechanical chores approach known as an alternative approach which sued for enhancement computer efficiency.

Although the straightforward alphabetic list approach for examining a target word is extra boring and time-consuming, authors proposed, and alternative methods called Lexical Matrix lacked. Moreover, the lexicalized concept works based on a lexical theory which represented some theory definitions of lexical semantics which relied on constructive or just differential.

An instructive method depends on the concepts, which provided either by human or machine. There is restricted to informative machine information due to the lack of data in some standard dictionaries.

Whereas, the WordNet Lexical database is strong relations due to using WordNet tool based on standard lexicography to show up the required data in front of the users. Finally, the authors developed the experimental study based on an on-line lexical dataset.

Furthermore, WordNet tool led to separate the dataset into two separate tasks: a lexical subfield of WordNet that is reported foundation documentation. On the other hand, initiate machine models.

As a result, they provided information about semantics which is more critical and complex to trim a word to its morphological features as a verb, noun, adjective, adverb, conjunction, etc.

Authors in this study [78] applied two methods for prepositional phrase attachment disambiguation. On the one hand, the selecting preference approach that decided proper candidate corresponding to an appropriate preposition goes like [VP, NP1]. On the other side, demonstration a vector space based on contextual identification words which represented an NP2.

Finally, they found out a strong association relation between verb phrase or first noun phrase and prepositions based on a specific criterion. As a result, they came up with three models: co-occurrence model, distributional semantic model, and word space model. Also, all these models rely on graph-based decency and algorithm of frequency count as 'log.'

In this study [79], authors propose unsupervised preposition sense disambiguation (PSD) system. That presented and compared to the usefulness of different methods and unsupervised training. Furthermore, they showed ways to extend this task to the degree of prepositional arguments. Moreover, a preposition acts as a link between two words the first is a noun head and an object.

Furthermore, the preposition senses can be accurately disambiguated using only the headword and object of the PP-attachment based on a study [80]. Also, they provided information about both Hidden Markov Model and Maximum Entropy algorithm.

Based on research [81], authors undertook a standard first order. Whereas, the SemEval [82] task which contained 16k of training and 8k test of sentences. Also, a TPP defined senses for each of the 34 most usage prepositions which proposed by [83] acquired an accuracy of 9% sense for each preposition. As a result, an accurate performance of 56%.

In this study [84], the authors projected a maximum entropy model (ME), using Wall Street Journal (WSJ) corpus which extracted from Penn Treebank. Moreover, they used ME model to detect verb phrases based on etymological information rather than

outward semantic acquaintance.

What's more, also, they trained and combined word classes on both ME model and binary hierarchy which derived by mutual information clustering from the corpus. Finally, they found out an accurate performance of 81.6%.

Authors in this study[85], mainly focused on providing a cursory treatment of the Arabic prepositional strategy. Also, they located a metric for an appropriate parser with the choice to distinguish an ambiguous content to clarify and assess unique parser settings.

Moreover, they applied a Depth First Search (DFS) algorithm to each parse tree using an Arabic Treebank (ATB) for extracting the dataset. Also, they divided dataset ATB in the experimental study into training, development, and test with the size of 18818 trees, 2318 trees, and 2313 trees, respectively.

Finally, they set the dataset using the split based on study proposed by [63]. As a result, they found out based on extracted from sentences from ATB of length reach at 40 words an accurate performance of 92.61%.

Authors in this study [86] suggested disambiguation prepositional phrase attachments. Although they solved since the relative strength of association of the preposition with verbal and nominal heads, they estimated since distribution an automatically parsed corpus.

Moreover, learning from a large corpus which contained 13 million words and they looked at the log of the ratio probability of verb attach to the likelihood of noun attachment. That based on candidates and the preposition. Lastly, PP-attachment attached preposition phrase to a more probable candidate. Accordingly, solving prepositional phrase attachment ambiguity based on lexical associations.

As a result, they obtained from the lexical associations, the performance of an accurate of 80%, 80%, and 86% based on precision, recall, and human judges respectively.

Authors in the study [87] proposed and applied disambiguation prepositional phrase techniques in English. Moreover, they provided information about many types of research which solved an arbitrary and inconsistent PP-attachment ambiguity.

Meanwhile, in this study, they suggested techniques for processing the definitions based on a method standard online dictionary.

Finally, authors proposed an alternative as opposed to heuristic utilized a traditional master framework shell with an implicit derivation motor. In this way, they chose to get rid of backtracking and with consistent unification and to outline our own more straightforward though speedier control structures. Moreover, they conducted dataset, which extracted from two online dictionaries: 'Webster's Seventh New Collegiate (W7)' and 'Longman Dictionary of Contemporary English (LDOCE).' Finally, they got a 'Choose function' disambiguation system. Also, various types of ambiguous sentences function Choose received each sentence, which may get only one interpretation based on its possible validation.

4. Discussion

This paper thoroughly reviewed different classes of ambiguity and then described various approaches to tackle PP-attachment ambiguity problem. The main contribution of our work is to compile different approaches to PP-attachment ambiguity resolution and thereby to make it easily accessible for comparative analysis. The most current approaches involve the classical minimalist approach, dictionary-based approach, corpus-based approach, machine-learning based approach and psycho-linguistically motivated approach. The latter method includes how human process the language in resolving the ambiguity.

It can be argued that despite the plethora of work on PP-attachment ambiguity resolution, the problem is still challenging for many underdeveloped languages like Arabic and Urdu for example. There could have been many reasons for this lack of work in such communications, notable lack of resources like language corpora. However, with the recent advances in computer technology and internet, a significant amount of text is freely available on the Web, which makes it possible to apply corpus-based data-driven approaches for ambiguity resolution.

Another important question, which came to the fore during this study, is that the different approaches have been applied in isolation without paying attention to their combined impact. Therefore, it might be interesting to see whether the performance of an ambiguity resolution system can be improved by combining the various approaches. Of course, this combination of methods could unfold new challenges notably time/space complexity of the overall system.

5. Conclusion

This paper aimed at providing an in-depth survey of the common approaches used to resolve PP-attachment ambiguities in natural language processing. We started by providing a detailed overview of different types of ambiguities in natural language processing and attempted to chart out a comprehensive taxonomy of these different types of ambiguities.

Then we narrowed down to PP-attachment ambiguity by comparing different ambiguity resolution approaches. These approaches range from the classical right/left-association approach to modern corpus-based approach. This survey suggests that despite the availability of different methods to resolving PP-attachment ambiguity, the problem remains to find such an approach, which can match the human-level skills of resolving such ambiguities.

There are two methods employed in natural language processing concerning new approaches: the first technique called the rule-based method and the second called statistical approach. Moreover, a rule-based method is standing up for laws in advance often by language experts and, then, design programs relating to any area of natural languages by such laws. Therefore, this method doesn't require earlier machine learning. Whereas, the statistical approach requires the presence of data prepared in advance in such a way that the computer can learn the laws, so persons make these data often are not at the expert level like the previous method.

References

[1] Daffa, W., Alshahry, R., Khan, I. H. Corpus-Based Prediction of Coordination Ambiguity in Arabic, 6 (3) 85–92.

[2] Mithchel, M. B. C., Feranando, M. Corley. (1995). Exposure-Based Models of Human Parsing: Evidence for the I. H. Use of Coarse-Grained (Nonlexical) Statistical Records.

[3] Khan, I. H., Haleem, M. (2013). Managing Lexical Ambiguity in the Generation of Referring Expressions, no. July, p. 33–39.

[4] Berry, K., Erik, D. (2003). From Contract Drafting to Software Specification: Linguistic Sources of Ambiguity, *Autom. Softw. Eng.* 1997. p 1–80.

[5] Hong, J. F. (2015). Verb sense discovery in Mandarin Chinese-a corpus based knowledge-intensive approach, p. 1–249.

[6] Ldmd, L. (1897). Lexical ambiguity revisited: on homonymy and polysemy, p. 1047–1054, 1897.

[7] Clahsen, H. (2009). Linguistics An Introduction, Second Edition. 2009.

[8] Dwivedi, V. D. (2013). Interpreting quantifier scope ambiguity: Evidence of heuristic first, algorithmic second processing, *PLoS One*, 8 (11) p. 1–20.

[9] Friederici, A. D., Steinhauer, K., Mecklinger, A., Meyer, M. (1998). Working memory constraints on syntactic ambiguity resolution as revealed by electrical brain responses, *Biol. Psychol.*, 47 (3) 193–221.

[10] Gillon, B. S. (1990). Ambiguity, generality, and indeterminacy: Tests and definitions, *Synthese*, 85 (3) 391–416.

[11] Hirst, G. (1988). Semantic interpretation and ambiguity, *Artif. Intell.*, 34 (2) 131–177.

[12] Merlo, P., Ferrer. E. E. (2006). The Notion of Argument in Prepositional Phrase Attachment, *Comput. Linguist.*, 32 (3) 341–377.

[13] Kamsties, E., Berry, D. M., Paech, B. (2001). Detecting Ambiguities in Requirements Documents Using Inspections,2016 *Engineering*, July, p. 1–12.

[14] Potts, C. (2014). Pragmatics, *Oxford Handb. Comput. Linguist.*, 2, p 1–27.

[15] Khan, I. H., van Deemter, K., Ritchie, G (2008). Generation of Referring Expressions: Managing Structural Ambiguities, *In:*

Proc. 22nd Int. Conf. Comput. Linguist. (Coling 2008), no. August, p. 433–440, 2008.

[16] Ungureanu, V. (2006). Referential Ambiguity- A source of Various meaning effects. p. 5, 2006.

[17] Poesio, M., Artstein, R. (2010). Annotating Anaphoric Ambiguity, Massimo Poesio Ron Artstein-University Essex, p. 1–19.

[18] Armstrong, B. C., Plaut, D. C. (2016). *Semantic Ambiguity Effects in Lexical Processing : A Neural-Network Account Based on Semantic Settling Dynamics*, September, 2016.

[19] Kimball, J. (1973). Seven principles of surface structure parsing in natural language, *Cogn.* 2(1)15–47.

[20] Yarowsky, D. (1995). Unsupervised Word Sense Disambiguation Rivaling Supervised Methods, *In*: Proc. 33rd Annu. Meet. Assoc. Comput. Linguist., p. 189–196, 1995.

[21] Ye, P., Baldwin, T. (2007). MELB-YB: Preposition Sense Disambiguation Using Rich Semantic Features, *In*: Proc. Fourth Int. Work. Semant. Eval., no. June, p. 241–244, 2007.

[22] Ratnaparkhi, A. (1998). Maximum Entropy Models for Natural Language Ambiguity Resolution, p. 163.

[23] Hara, T. O. (2003). Preposition Semantic Classification via PENN TREEBANK and FRAMENET, 75, May, p 30–31.

[24] Bruce, R. F., Wiebe, J. M. (1999). Decomposable Modeling in Natural Language Processing, *Comput. Linguist.* 25 (2) June 1999.

[25] Marcus, M. P., Santorini, B., Marcinkiewicz, M. A. (1993). Building a large annotated corpus of English: The Penn Treebank., *Comput. Linguist.*, 19 (2) 313–330.

[26] Ratnaparkhi, A. (1998). Statistical Models for Unsupervised Prepositional Phrase.

[27] Harabagiu, S. M. (1996). An Application of WordNet to Prepositional Attachment, p. 360–362.

[28] Nal, C. H. (1991). Resolving a pragmatic prepositional phrase attachment ambiguity, p. 351–352.

[29] Stetina, J., Nagao, M., Honmachi, Y. (1997). Corpus based PP attachment ambiguity resolution with a semantic dictionary, *Proc. Fifth Work. Very Large*, p. 66–80.

[30] Litkowski, K. (2014). Pattern Dictionary of English Prepositions, *Acl*, p. 1274–1283.

[31] Cimiano, P., Minock, M (2010). Natural language interfaces: what is the problem?–a data-driven quantitative analysis, *In*: *Proc. 14th Int. Conf. Appl. Nat. Lang. to Inf. Syst.*, p. 192–206, 2010.

[32] Decker, S., Erdmann, M., Fensel, D., Studer, R. (1999). Ontobroker: Ontology based access to distributed and semi-structured information, *Database Semant. IFIP — Int. Fed. Inf. Process.* 11, Gruber 93, p. 351–369.

[33] Alam, Y. S. (2004). Decision Trees for Sense Disambiguation of Prepositions: Case of Over, *HLT-NAACL 2004 Work. Comput. Lex. Semant.*, p 52–59.

[34] Green, S. (2009). Improving Parsing Performance for Arabic PP Attachment Ambiguity.

[35] Huyck, C. R. (2002). Prepositional Phrase Attachment Ambiguity Resolution Using Semantic Hierarchies, *Science (80-.)* p 1–4.

[36] Amherst, M. (1988). Itegration of Semantic and Synatactic Constraints for Structural Noun Phrase Disambiguation, p. 1486–1491.

[37] Maamouri, M., Bies, A., Kulick, S. (2008). Enhanced annotation and parsing of the arabic treebank, *Proc. INFOS*, p 176–183.

[38] Diab, M. T. (2007). mproved Arabic Base Phrase Chunking with a new enriched POS tag set, *Comput. Linguist.*, no. June, p. 89–96, 2007.

[39] Chantree, F., Willis, A., Kilgarriff, A., DeRoeck, A. (2007). Detecting dangerous coordination ambiguities using word distribution, *Recent Adv. Nat. Lang. Process.* 4, p. 287–296.

[40] Kilgarriff, A., Rychly, P., Smrz, P., Tugwell, D. (2004). The Sketch Engine, *Proc. Elev. EURALEX Int. Congr.*, p. 105–116.

[41] Nadh, K., Huyck, C. R. (2002). Prepositional Phrase Attachment Ambiguity Resolution Using Semantic Hierarchies, *Science* (80-.) p. 1–4.

[42] Nakov, P., Hearst, M. (2005). Using the web as an implicit training set: Application to structural ambiguity resolution, HLT/ EMNLP 2005 - Hum. Lang. Technol. Conf. Conf. Empir. Methods Nat. Lang. Process. Proc. Conf., no. October, p. 835–842.

[43] Al-sabbagh, R. (2008). A Web-Based Approach for Arabic PP Attachment, p. 22–26.

[44] Pantel, P., Lin, D. (2000). An unsupervised approach to prepositional phrase attachment using contextually similar words, *In:* Proc. 38th Annu. Meet. Assoc. Comput. Linguist., p 101–108.

[45] Warren, P., Schafer, A. J., Speer, S. R., White, S. D. (2000). Prosodic Resolution of Prepositional Phrase Ambiguity in Ambiguous and Unambiguous Situations, *Univ. Calif. Work. Pap. Phonetics*, 99, December p 5–33, 2000.

[46] Rakshit, G., Sontakke, S., Bhattacharyya, P., Haffari, G. (2016). Prepositional Attachment Disambiguation Using Bilingual Parsing and Alignments.

[47] Bailey, D., Lierler, Y., Susman, B., Prepositional Phrase Attachment Problem Revisited: How, no. October, p. 12–22, 2015.

[48] Al-sabbagh, R., El-zeiny, N. (2007). Arabic Anaphora Resolution Using the Web as Corpus, no. December, p 5–6.

[49] Spivey-Knowlton, M., Sedivy, J. C. (1995). Resolving attachment ambiguities with multiple constraints, *Cognition*, vol. 55, (3) 227–267.

[50] Cedex, O. (2003). Lexicalising a robust parser grammar using the WWW, p. 620–626.

[51] Aït-Mokhtar, S., Chanod, J.-P., Roux, C. (2002). Robustness beyond shallowness: incremental deep parsing, *Nat. Lang. Eng.*, 8 (2–3) 0–24.

[52] Ratnaparkhi, A. (1998). Statistical models for unsupervised prepositional phrase attachment, *In:* Proc. 36th Annu. Meet. Assoc. Comput. Linguist. -, 2, p 1079–1085.

[53] Merlo, P., Crocker, M., Berthouzoz, C. (1997). Attaching Multiple Prepositional Phrases: Generalized Backed-off Estimation, *In: Proc. Second Conf. Empir. Methods Nat. Lang. Process.*, p. 145–154, 1997.

[54] Collins, M., Brooks, J. (1995). Prepositional Phrase Attachment through a Backed-Off Model, *ACL WS Very Large Corpora*, p. 12.

[55] Olteanu, M., Moldovan, D. (2014). PP-attachment disambiguation using large context, *Uma ética para quantos?*, vol. XXXIII, no. October, p. 81–87.

[56] Greenberg, C. (2014). Disambiguating prepositional phrase attachment sites with sense information captured in contextualized distributional data, *In:* Proceedings ACL 2014 Student Res. Work., p 71–77, 2014.

[57] Cohen, J. (1960). A coefficient of agreement for nominal scales, XX, (1) 37–46, 1960.

[58] Srikumar, V., Roth, D. (2013). Modeling semantic relations expressed by prepositions, *Trans. Assoc. Comput. Linguist.*, . 1, no. 2013, p. 231–242.

[59] Rah, A., Adone, D. (2010). L2 Learners' Processing of PP Attachment Ambiguities – A Production Study, *Second Lang. Res.*, p. 139–149.

[60] Aguilar, N., Alonso Alemany, L., Lloberes, M., Castellón, I. (2011). Resolving prepositional phrase attachment ambiguities in Spanish with a classifier, *Proces. Leng. Nat.*, 46, p. 75–82.

[61] Atterer, M., Schütze, H. (2007). Prepositional Phrase Attachment without Oracles, *Comput. Linguist.*, 33 (4) 469–476.

[62] Bikel, D. M. (2004). Intricacies of Collins' Parsing Model, *Comput. Linguist.*, 30 (4) 479–511.

[63] Katsika, K. (2011). Attachment preferences and corpus frequencies in PP ambiguities: Evidence from Greek 1 Kalliopi Katsika.

[64] Coppola, G. F., Birch, A., Deoskar, T., Steedman, M. (2011). Simple Semi-Supervised Learning for Prepositional Phrase Attachment, *In: Proc. 12th Int. Conf. Parsing Technol.*, p. 29–139.

[65] Litkowski., Hargraves, O (2007). SemEval-2007 Task 06: Word-Sense Disambiguation of Preposition, *In:* Proc. 4th Int. Work. Semant. Eval., no. June, p 24–29.

[66] Pulman, S. G. (1991). Basic Parsing Techniques : an introductory survey, *Univ. Cambridge Comput. Lab. SRI Int.*, no. April, p. 1–20.

[67] Shaalan, K., Hendam, A., Rafea, A. (2010). An English-Arabic bi-directional machine translation tool in the agriculture domain: A rule-based transfer approach for translating expert systems, *IFIP Adv. Inf. Commun. Technol.*, vol. 340 AICT, p. 281–290.

[68] Houpt, J. (2006). Ambiguous Prepositional Phrase Resolution by Humans, *Citeseer*, 2006.

[69] Huyck, C. R. (1998). A Practical System for Human-Like Parsing.

[70] Brill, E., Resnik, P. (1994). A rule-based approach to prepositional phrase attachment disambiguation, *In:* Proc. 15th Conf. Comput. Linguist. 2, p. 1198–1204.

[71] Zhao, S., Lin, D. (2003). A Nearest-Neighbor Method for Resolving PP-Attachment Ambiguity, 8, Lin.

[72] Belinkov, Y. (2014). Neural Network Architectures for Prepositional Phrase Attachment Disambiguation.

[73] Levesque, H. J. (2011). The Winograd Schema Challenge, *Artif. Intell.*, 1989, p. 63–68.

[74] Olteanu, M., Moldovan, D. (2005). PP-attachment disambiguation using large context, *In:* Proc. Conf. Hum. no. October, p. 273–280.

[75] Adwait Ratnaparkhi. (1996). A maximum entropy model for part-of-speech tagging, *In:* Proc. Empir. Methods Nat. Lang. Process. Conf., 1 (49)133–142.

[76] Pan, H. Y., Felser, C. (2010). Referential context effects in L2 ambiguity resolution: Evidence from self-paced reading, *Lingua*.

[77] Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D., Miller, K. J. (1990). Introduction to wordnet: An on-line lexical database, *Int. J. Lexicogr.*, 3 (4) 235–244.

[78] Simon Suster, C. C. (2012). European master in Language and Communication Technologies Université de Lorraine & Rijksuniversiteit Groningen.

[79] Hovy, D., Vaswani, A., Tratz, S., Chiang, D., Hovy, E., Way, A., Rey, M. (2011). Models and Training for Unsupervised Preposition Sense Disambiguation, *Comput. Linguist.*, p. 323–328.

[80] Hovy, D., Tratz, S., Hovy, E. (2010). What's in a Preposition? Dimensions of Sense Disambiguation for an Interesting Word Class, *Coling*, no. August, p. 454–462, 2010.

[81] Dempster, A. P., Laird, N. M., Rubin, D. B. (1977). Maximum likelihood from incomplete data via the {EM} algorithm, *J. R. Stat. Soc. Ser. B*, 39 (1) 1–38.

[82] J. G. and S. S. J. Artiles. (2007). The SemEval-2007 WePS Evaluation: Establishing a benchmark for the Web People Search Task, *Production*.

[83] Baldwin, T., Kordoni, V., Villavicencio, A. (2009). Prepositions in Applications: A Survey and Introduction to the Special Issue, *Comput. Linguist.*, 35 (2) 119–149.

[84] Ratnaparkhi, A., Reynar, J., Roukos, S. (1994). A maximum entropy model for prepositional phrase attachment, *Proc. Work. Hum. Lang. Technol.*, p. 250–255.

[85] Chiang, D., Diab, M., Habash, N., Rambow, O., Shareef, S. (2005). Parsing Arabic Dialects, *Final Report, 2005 JHU Summer Work.*, p. 369–376, 2005.

[86] T. B. Laboratories and Hill, M. (1990). Structural Ambiguity and Lexical Relations NJ 07974, p. 229–236.

[87] Jensen, K., Binot, J.-L. (1987). Disambiguating Prepositional Phrase Attachments by Using On-line Dictionary Definitions, *Comput. Linguist.*, 13 (3–4) 251–260.